Framework for the evaluation of

The Smith Family's

Learning for Life scholarship program

November 2006

THE SMITH FAMILY

everyone's family

Note

This framework is closely modelled on the framework developed by the *Centers for Disease Control and Prevention* (CDC) in the USA.[1] It has the same structure and contains extensive passages taken verbatim from the CDC document. It also draws material extensively from related documents prepared by the CDC, especially a document titled: *Introduction to program evaluation for public health programs: a self-study guide*.[2] This document contains many checklists and worksheets that were adapted to suit the *Learning for Life* evaluation framework.

As the *Learning for Life* evaluation framework document is an adaptation of the CDC framework, there is no author for it. However, responsibility for the contents of the *Learning for Life* framework document is taken by Adrian Beavis and Maree Murray of the Research and Evaluation Group, within The Smith Family, Australia. The Framework has been approved by The Smith Family's Strategy Group. We wish to acknowledge the encouragement received from staff at the CDC in the preparation of this document.[3] We also wish to thank the members of The Smith Family's Research Advisory Committee for reviewing the document. We alone, however, are responsible for any errors in this document.

Adrian Beavis and Maree Murray

The Smith Family

November 2006

---

1 Centers for Disease Control and Prevention (1999). Framework for Program Evaluation in Public Health MMWR 1999; 48 (No. RR-11), available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr4811a1.htm, viewed May 2006.

2 http://www.cdc.gov/eval/evalguide_pt01.pdf, viewed May 2006.

3 The style of this document is based upon the *Publication Manual of the APA, 5th Edition*, the American Psychological Association, as implemented by the bibliographic computer program *EndNote* (*EndNote* does not include the full spectrum of reference types that are described in the APA manual).

Table of contents

Checklists

Worksheets

Overview of the framework for the evaluation of The Smith Family's Learning for Life scholarship program

Research seeks to prove; evaluation seeks to improve (M. Q. Patton)

The *Learning for Life* scholarship program is an education support program that provides financial assistance to help families with the education needs of their children and personal support delivered through *Learning for Life* Workers. *Learning for Life* Workers provide advice and knowledge concerning other educational, health and social support services, and access to other appropriate Smith Family programs in the *Learning for Life* suite.

The objective of this framework is to facilitate an effective evaluation of *Learning for Life.*[4]

Effective program evaluation is a systematic way to improve and develop a program using procedures that are useful, feasible, ethical, and accurate.

This framework describes these procedures and provides reasons for their use. It describes the tasks associated with these procedures. Checklists and worksheets are provided to help implement these procedures. The framework also provides options for stakeholders to consider at important points of the evaluation.

The framework is organised around the six key tasks or steps that are required for the evaluation to be effective. These are:

Step 1: engage stakeholders

Step 2: describe the program

Step 3: focus the evaluation design

Step 4: gather credible evidence from a variety of sources

Step 5: justify conclusions

Step 6: ensure use.

The framework also includes a timeline for the evaluation.

---

4 While acknowledging the need for evaluations of the broad suite of *Learning for Life* programs, the centrality and extent of the scholarship program and the need for a strong, centred evaluation means that this framework has been drafted to focus solely on the scholarship program of *Learning for Life*. Throughout this document, then, unless otherwise specified, *Learning for Life* is taken to mean the *Learning for Life* scholarship program.

1. This framework document is a blueprint for the conduct of the evaluation.	This document is the first major step in the evaluation of the *Learning for Life* program. It describes at a high level of generality what, why, how and with whom this evaluation is to be undertaken. By analogy, it is akin to a blueprint for the conduct of the project. It also provides a timeline for the evaluation.
2. The framework is based upon the one prepared by the CDC.	This document is modelled upon the framework developed by the *Centers for Disease Control and Prevention* (1999) (CDC) in the USA[5] and draws material extensively from a related document also prepared by the CDC.[6]
3. The CDC framework is well suited to the *Learning for Life* evaluation.	Although the CDC has a focus upon community health, their framework was readily adaptable to the evaluation of *Learning for Life*. As the CDC document notes: By integrating the principles of this framework … we will stimulate innovation toward outcome improvement and be better positioned to detect program effects. More efficient and timely detection of these effects will enhance our ability to translate findings into practice. … Findings … will lead to program plans that are clearer and more logical; stronger partnerships will allow collaborators to focus on achieving common goals; integrated information systems will support more systematic measurement; and lessons learned from evaluations will be used more effectively to guide changes in … strategies.[7]

5 http://www.cdc.gov/mmwr/preview/mmwrhtml/rr4811a1.htm, viewed May 2006.

6 U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. Office of the Director. Office of Strategy and Innovation. 2005. *Introduction to program evaluation for public health programs: A self-study guide*. Atlanta, GA: Centers for Disease Control and Prevention.

7 'Forward' by Koplan, J. P., Director, Centers for Disease Control and Prevention Administrator, Agency for Toxic Substances and Disease Registry.

	The orientation and planned outcomes derived from the CDC framework coincides well with the evidence-based, collaborative approach espoused by The Smith Family. As the CDC (1999) notes: > The emphasis is on practical, ongoing evaluation strategies that involve all program stakeholders, not just evaluation experts. Understanding and applying the elements of this framework can be a driving force for planning effective … strategies, improving existing programs, and demonstrating the results of resource investments.
4. Main questions addressed by the framework.	The framework addresses three main questions: 1. What is the best way to evaluate the program? 2. What is being learned from the evaluation? 3. How should the evaluation be used to further develop the program? This framework helps answer these questions by suggesting strategies that are useful, appropriate, feasible and ethical. When applying the framework, the challenge is to devise an optimal – as opposed to an ideal – strategy. An optimal strategy is one that accomplishes each step in the framework in a way that accommodates the program context and meets, or exceeds, all relevant standards.
5. There are three points of reference in making judgements about *Learning for Life*.	The evaluation of *Learning for Life* – as with any evaluation of a social program – involves making judgements about the program, referencing its: 1. merit (that is, quality) 2. worth (that is, cost-effectiveness) 3. significance (that is, importance). A program may have merit, but not be cost effective or significant. It may be worthwhile, but of low importance. Judgements such as these are based upon values.

6. There are six main questions that need to be addressed to evaluate *Learning for Life*. This framework provides procedures for answering these questions.	Judgements regarding a program imply the following questions have been answered: 1. What will be evaluated? (That is, what precisely is the *Learning for Life* scholarship program and in what context does it exist?) 2. What aspects of *Learning for Life* should be considered when judging program performance? 3. What standards (that is, type or level of performance) must be reached for *Learning for Life* to be considered successful? And related to this: What measures can indicate these standards? 4. What evidence will be used to indicate how *Learning for Life* has performed? 5. What conclusions regarding program performance are justified by comparing the available evidence to the selected standards? 6. How will the lessons from the evaluation be used to improve and develop *Learning for Life*? This framework for evaluating *Learning for Life* is designed to help answer these six questions *as part of the evaluation process*.
7. To answer the six evaluation questions, there are six steps which need to be taken.	To answer the six evaluation questions, there are six steps which need to be taken. The framework sets out the six steps that must be taken. They are starting points for tailoring the evaluation to the particular context of *Learning for Life*. Earlier steps provide the foundation for subsequent ones. The steps are: Step 1: engage stakeholders Step 2: describe the program Step 3: focus the evaluation design Step 4: gather credible evidence

	Step 5: justify conclusions Step 6: ensure use and share lessons learned. Adhering to these steps in this order will facilitate a deeper understanding of *Learning for Life*'s history, setting and organisation. This, in turn, should ensure that the evaluation is well conceived and conducted and well accepted by stakeholders.
8. The framework also provides standards for use in the evaluation of *Learning for Life*.	The framework also provides a set of standards for assessing the quality of evaluation activities. These standards are organised into the following four groups: Standard 1: utility Standard 2: feasibility Standard 3: propriety Standard 4: accuracy. These standards help answer the question, 'Is it likely that this evaluation will be effective?' and are planned as criteria for judging the quality of the evaluation efforts.

9. Common concerns with evaluation are addressed by this framework.	Common concerns regarding program evaluation may also be clarified by using this framework. These concerns include: 1. the evaluation will be too costly. The expense of an evaluation is relative; the cost depends on the questions being asked and the level of precision desired for their answers. A simple, low-cost evaluation can deliver valuable results. 2. the evaluation is time-consuming and tangential to program operations. This framework encourages conducting an evaluation that is timed strategically to provide necessary feedback to guide action. 3. the perceived technical demands of designing and conducting the evaluation are too high. Although circumstances exist where controlled environments and elaborate analytic techniques are needed, most evaluations do not require such methods. Instead, the practical approach endorsed by this framework focuses on questions that seek to develop *Learning for Life* by using context-sensitive methods and analytic techniques that summarise accurately the meanings of qualitative and quantitative information encountered in the evaluation. 4. the evaluation will trouble some program staff because they may perceive it as punitive, exclusionary, or adversarial. This framework encourages an evaluation approach that is designed to be helpful and engages all interested stakeholders in processes that encourage their participation.

10. Selecting an evaluation team. *See Appendix A for details.*	Appendix A provides information to assist in selecting an evaluation team. It recommends a team with one person as the evaluation leader to conduct the evaluation and having at least some of that team located within The Smith Family.
11. Risk and risk management.	In this chapter and each of the following, a part is devoted to identifying potential risks and possible responses for the management of them. In this Overview chapter, two global risks are identified. In other sections, risks peculiar to specific elements of the evaluation are described. The following are seen as potential global risks to the evaluation. The first is seen as especially important. 1. <u>Risk</u>: the evaluation cannot establish any positive outcomes or, it establishes that *Learning for Life* is low in merit, worth or significance, leading to the conclusion that the program ought to be abandoned. <u>Management of risk:</u> The Smith Family management to develop a set of policy options to respond to this outcome before Step 1 commences, so that lessons learned from the evaluation can be used, and improvements to the program can be efficiently implemented, or the development of a replacement program is appropriately organised as this outcome begins to seem increasingly likely. 2. <u>Risk</u>: the principal evaluator or other key members of the evaluation team leave the project. <u>Management of risk</u>: ensure that the evaluation team works collaboratively distributing information freely, that the team has a strategy in place to identify replacements if necessary, and that the names of potential successors to replace key personnel are noted and regularly updated at all times during the evaluation.

12. Products of the evaluation of *Learning for Life.*	Appendix E lists the nine main products of the evaluation of *Learning for Life*. A main report, a summary report and a technical report are the major products.
13. Next steps.	The remainder of this framework discusses each of the major six steps and how the standards that govern effective program evaluation apply. Following each step is an outline of risks and matched risk management tactics. Support materials in the form of checklists and worksheets are included and the timeline or the evaluation are also attached.

Step 1: Engaging stakeholders

> *The underlying logic … is that good evaluation does not merely gather accurate evidence and draw valid conclusions, but produces results that are used to make a difference. To maximize the chances evaluation results will be used, you need to create a "market" before you create the "product"—the evaluation. You determine the market by focusing your evaluations on questions that are most salient, relevant, and important.*[8]

An evaluation begins by engaging stakeholders – the persons or organisations having an investment in what will be learned from the evaluation and what will be done with the knowledge.

14. Engaging stakeholders in the evaluation is fundamental to the success of the evaluation.	Stakeholders must be engaged in the inquiry to ensure their perspectives are recognised. If stakeholders are not engaged, the evaluation might not address important elements of a program's objectives, operations and outcomes. Therefore, evaluation findings might be ignored, criticised or resisted because the evaluation did not address the stakeholders' concerns, values or understandings. Stakeholders may help in the other steps.
15. There are three main groups of stakeholders.	Identifying and engaging the following three principal groups of stakeholders will be critical for the evaluation of *Learning for Life*: 1. **those involved** in program operations (for example, sponsors, collaborators, partners, administrators, managers, and staff) 2. **those served** or affected by the program (for example, *Learning for Life* scholarship holders (present, and if possible, past), family members, neighbourhood organisations, sceptics, and staff of related or competing organisations) 3. **primary users** of the evaluation.

[8] U.S. Department of Health and Human Services. CDC. Office of the Director. Office of Strategy and Innovation 2005, p. 6.

16. Those involved in the operation of *Learning for Life* as stakeholders may not be a single interest group. They should not confuse the program evaluation with personnel appraisal.	**Those involved in *Learning for Life*** Persons or organisations involved in *Learning for Life* operations have a stake in how the evaluation activities are conducted because *Learning for Life* might be altered as a result of what is learned. Although staff work together on *Learning for Life*, they are not necessarily a single interest group. Subgroups might hold different perspectives and follow alternative agendas. Further, because these stakeholders have a professional role in the program, they might perceive the evaluation as an effort to judge them personally. Program evaluation is related to, but must be distinguished from, individual performance appraisal.
17. Those served or affected by *Learning for Life* may need to include critics of the program.	**Those affected by *Learning for Life*** Persons or organisations affected by *Learning for Life*, either directly (for example, by receiving support) or indirectly will be identified and engaged in the evaluation to the extent possible. Although engagement of supporters of *Learning for Life* is expected, individuals who are sceptical or critical of the program are also important to engage. If these exist, they will need to be identified. (Disagreement about *Learning for Life* might stem from differing values regarding what change is needed or how to achieve it.) Opening the evaluation to different perspectives and enlisting the help of critics in the inquiry is prudent because these efforts can strengthen the evaluation's credibility, and ultimately *Learning for Life* itself and The Smith Family.

18. Primary users of the evaluation must be engaged early on and their involvement maintained.	**Primary users of the evaluation** Primary users of the evaluation are the specific persons who are in a position to do, or decide, something major regarding the conduct of *Learning for Life*. In practice, primary users will be a subset of all stakeholders identified. To be successful the evaluation will need to designate primary users early in its development and maintain frequent interactions with them so the evaluation addresses their values and satisfies their unique needs.
19. The level of involvement may vary for different stakeholders.	The scope and level of involvement may vary for each stakeholder in the evaluation. Some stakeholders can be directly involved in designing and conducting the evaluation, others can be kept informed regarding progress of the evaluation through periodic meetings, reports and other means of communication. Trust among stakeholders is essential; therefore caution is required to prevent misuse of the evaluation process. Similarly, it is imperative that strong relations, communication and understandings are established with stakeholders from the outset and subsequently maintained.
20. Standards for engaging stakeholders.	Each of the four standards (see paragraph number 8 in the previous chapter) give rise to a specific question that must be addressed when considering the engagement of stakeholders: 1. Utility: Who will use these results? 2. Feasibility: How much time and effort can be given to stakeholder engagement? 3. Propriety: Which stakeholders need to be consulted to ensure the evaluation is conducted ethically? 4. Accuracy: How should stakeholders be engaged to ensure an accurate account of *Learning for Life*?

21. Questions to be asked of all stakeholders.	Stakeholders in *Learning for Life* will be asked the following questions: 1. Who do you represent? 2. Why are you interested in *Learning for Life*? 3. What is important about *Learning for Life* to you? 4. What do you think *Learning for Life* accomplishes? 5. What would you like *Learning for Life* to accomplish? 6. How much progress would you expect *Learning for Life* to have made by now? 7. What do you currently see as the critical evaluation questions? 8. How do you think you will use the results of this evaluation? 9. What resources (that is, time, funds, knowledge, evaluation expertise, access to respondents, and access to policymakers) could you contribute to this evaluation effort?
22. Risks and risk management.	The following are seen as potential risks to Step 1: 3. Risk: stakeholders do not engage with the evaluation process. Management of risk: senior management to be proactive in their support of the evaluation. Opportunities for managed engagement are available throughout the evaluation period to foster strong, productive engagement. 4. Risk: insufficient stakeholders are able to be engaged, or as a group, not usefully representative of *Learning for Life* stakeholders. In particular there may be difficulties engaging previous *Learning for Life* participants who may be disenchanted with aspects of the program. Management of risk: stringent efforts are undertaken (and documented) by the evaluators, to engage a useful diversity

	of stakeholders. Engagement is seen as an ongoing, managed process, so that stakeholders who were initially absent, reluctant, or slow to engage, are able to contribute in a useful fashion as the evaluation proceeds.
23.Next steps.	On the following pages there is: • a checklist of activities that need to be undertaken to engage stakeholders • a set of worksheets designed to facilitate identifying and working with stakeholders. Completion of checklists and worksheets is planned as part of the work of the evaluation.

Checklist 1: Engaging stakeholders

- [ ] Address each of the four questions set by the standards for engaging stakeholders (utility, feasibility, propriety, accuracy).

- [ ] Identify stakeholders, using the three broad categories: (1) those affected, (2) those involved in operations, and (3) those who will use the evaluation results. If stakeholders fit more than one category they should be placed in all relevant categories.

  (Indicate if a particular stakeholder is a critic of *Learning for Life*.)

- [ ] Review the initial list of stakeholders to identify key stakeholders needed to improve credibility, implementation, advocacy, resourcing and authorisation decisions in the evaluation.

- [ ] Engage individual stakeholders or representatives of stakeholder organisations, or both.

- [ ] Create a plan for stakeholder involvement and identify areas for stakeholder input.

- [ ] Target selected stakeholders for useful participation in key steps, including writing the program description, suggesting evaluation questions, choosing evaluation questions, and disseminating evaluation results.

Worksheet 1: Identifying key stakeholders in Learning for Life

	Stakeholder is needed to:			
Indicate if any stakeholders are a critic of Learning for Life *by highlighting them in the list below.*	Increase credibility	Implement interventions for evaluation	Advocate for change	Funding for continuation
Stakeholders involved in the operations of *Learning for Life*				
1.	☐	☐	☐	☐
2.	☐	☐	☐	☐
3.	☐	☐	☐	☐
4.	☐	☐	☐	☐
5.	☐	☐	☐	☐
6.	☐	☐	☐	☐
Stakeholders affected by *Learning for Life*				
1.	☐	☐	☐	☐
2.	☐	☐	☐	☐
3.	☐	☐	☐	☐
4.	☐	☐	☐	☐
5.	☐	☐	☐	☐
6.	☐	☐	☐	☐
Stakeholders who will use the evaluation results				
1.	☐	☐	☐	☐
2.	☐	☐	☐	☐
3.	☐	☐	☐	☐
4.	☐	☐	☐	☐
5.	☐	☐	☐	☐
6.	☐	☐	☐	☐

Worksheet 2: What matters most to the key stakeholders in Learning for Life

Stakeholder	**What activities or outcomes of *Learning for Life* matter most to this stakeholder?**
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	

Step 2: Describing Learning for Life

A description of *Learning for Life* should convey its mission and those of its objectives being evaluated. This description should:

1. be sufficiently detailed to ensure there is a good understanding of goals and strategies of *Learning for Life*
2. discuss *Learning for Life*'s capacity to effect change, its stages of development, and how it fits with the Smith Family's vision and strategies, and with the wider community
3. set the frame of reference for all subsequent decisions in the evaluation
4. enable comparisons with similar programs (if these exist)
5. facilitate attempts to connect *Learning for Life*'s components to their effects.

Stakeholders might have differing ideas regarding the goals and purposes of *Learning for Life*. If the evaluation is conducted without agreement on what constitutes *Learning for Life* then it is likely to be of very limited use.

Aspects for inclusion in a description of *Learning for Life* are:

1. need
2. expected effects
3. activities
4. resources
5. stages of development
6. context
7. logic model.

24. Statement of need – the problems that *Learning for Life* addresses.	A statement of need describes the problems that *Learning for Life* addresses and implies how it responds. Important features for describing need include: 1. the nature and magnitude of the problems 2. which groups are affected by these problems 3. whether the need is changing and, if so, in what manner the need is changing.
25. Expected effects - what *Learning for Life* must accomplish to be considered successful.	A description of expectations conveys what *Learning for Life* must accomplish to be considered successful. As *Learning for Life*'s effects will unfold over time, the description of expectations needs to be organised chronologically. These will range from short-term to long-term consequences. *Learning for Life*'s mission, goals and objectives represent varying levels of specificity regarding expectations.
26. Activities – what *Learning for Life* does to create change.	Describing the activities of *Learning for Life* lays out what it does to facilitate change. By linking specific steps to strategies and actions in a logical sequence, *Learning for Life*'s mechanisms for change are also described. The description of activities distinguishes activities that are the direct responsibility of *Learning for Life* from those that are conducted by other programs or partners. External factors that might influence *Learning for Life*'s success should also be noted.

27. Resources – what is available for *Learning for Life* to conduct its activities.	Resources available to *Learning for Life* to conduct its activities include: 1. time 2. expertise 3. technology and other equipment 4. information 5. money, financial support and other assets available 6. goodwill. The description of *Learning for Life*'s resources should convey both their amount and intensity, highlighting where mismatches exist between desired activities and resources available to execute those activities.
28. The stage of development of *Learning for Life* shapes the evaluation questions.	The changing maturity of *Learning for Life* should be considered during the evaluation. Three broad stages of development can be identified in the life of projects: 1. Planning: during planning, program activities are untested, and the goal of evaluation is to refine plans 2. Implementation: during implementation, program activities are being field-tested and modified. At this stage the goal of evaluation is to characterise real, as opposed to ideal, program activities and to improve operations, perhaps by revising plans 3. Effects: during the last stage, enough time has passed for the program's effects to emerge; the goal of evaluation is to identify and account for both intended and unintended effects. It is necessary to identify the stage (or stages of different parts) at which *Learning for Life* is operating.

29. The context in which *Learning for Life* operates has to be understood for a good evaluation and to interpret findings.	Descriptions of *Learning for Life*'s context include its setting and particular environmental influences under which it operates (for example, history, geography, politics, social and economic conditions, and efforts of related or competing organisations). Understanding these environmental influences is required to: 1. design a context-sensitive evaluation 2. aid users in interpreting findings accurately and assessing the generalisability of the findings.
30. Logic models will be developed for the evaluation of *Learning for Life*. Appendix B provides a guide to this task.	Logic models are graphical depictions of the relationship between a program's activities and its intended outcomes. The logic model for the evaluation describes the sequence of events for bringing about change by synthesising the main elements of *Learning for Life* into a picture of how it is supposed to work. This model is displayed as a flow chart, to portray the sequence of steps leading to the results it has produced. Families of logic models may be created to display different levels of detail, from different perspectives, or for different audiences. One of the virtues of the logic model is its potential to summarise *Learning for Life*'s overall mechanism of change by linking processes to outcomes. The logic model can also display the infrastructure needed for *Learning for Life*'s operations. Elements that are connected within a logic model generally include inputs, outputs and results. These range from immediate- to intermediate- to long-term effects. Creating a logic model can help stakeholders to clarify *Learning for Life*'s strategies. A detailed logic model may also strengthen claims of causality and be a basis for estimating the effects of *Learning for Life*. Appendix B provides a guide to constructing a logic model.

31. Checking the accuracy of the description of *Learning for Life*.	The accuracy of the description of *Learning for Life* should be confirmed through consulting with its stakeholders. Where possible reported descriptions of its practice should be checked against direct observation of activities in the field.
32. Standards for describing *Learning for Life*.	Each of the four standards give rise to questions which need to be addressed when describing the *Learning for Life* program: 1. Utility: a. Is the level of detail appropriate or is there too much or too little detail? b. Is the description intelligible to those who need to use it? 2. Feasibility: a. Does the description include at least some activities and outcomes that are in control of *Learning for Life*? 3. Propriety: a. Is the evaluation plan comprehensive and balanced in assessing key aspects of *Learning for Life*, including its strengths and weaknesses? b. Does the description include sufficient detail to examine both strengths and weaknesses, and unintended as well as intended outcomes? 4. Accuracy: a. Is the description of *Learning for Life* comprehensive? b. Has the context of *Learning for Life* been documented so that likely influences on it can be identified?

33. Risks and risk management.	The following are seen as potential risks to Step 2: 5. <u>Risk</u>: differing stakeholder perspectives prevent one coherent description of *Learning for Life* being established. <u>Management of risk</u>: if there are disparate and divergent views these will be noted and considered throughout the course of the evaluation by the evaluators. For clarity however, a prime view and description of *Learning for Life* must be established during Step 2 of the evaluation. 6. <u>Risk</u>: suitable comparator programs may not be located or accessible. <u>Management of risk</u>: the evaluators will document search efforts for suitable comparator programs. 7. <u>Risk</u>: attempts to connect *Learning for Life*'s components to their effects are too general or inaccurate as the knowledge base at this stage of the evaluation is not strong. <u>Management of risk</u>: the evaluators will document these limitations in information and record how and why they have occurred. As the evaluation proceeds, understandings of the connections between *Learning for Life* components and effects may be improved.
34. Next steps.	On the following page there is: • a checklist of activities that need to be undertaken to describe *Learning for Life* in terms suited to the conduct of a sound evaluation. A set of worksheets designed to facilitate the development of a logic model of *Learning for Life* accompanies Appendix B. This appendix provides a guide to developing a logic model. The next chapter, *Step 3: Focusing the Evaluation Design*, shows how to use a logic model to identify and address issues of evaluation focus and design. Before this, it will be necessary to have developed the logic model.

Checklist 2: Describing Learning for Life

- [ ] Compile a comprehensive program description including needs, targets, outcomes, activities and resources.
- [ ] Identify stages of development and the program context.
- [ ] Convert inputs, activities, outputs and outcomes into a simple global logic model.
- [ ] Develop more detailed models from the global model as needed.

Step 3: Focusing the evaluation design

The evaluation should be focused to assess and include the issues of greatest concern to stakeholders while using time and resources as efficiently as possible.

Not all design options are equally well-suited to meeting the needs of stakeholders.

Among the items to consider when focusing the evaluation are its purpose, users, uses, questions, methods and agreements.

35. Purposes of an evaluation may include:

- *Gaining insight*
- *Changing practice*
- *Assessing effects*
- *Affecting participants.*

Characteristics of *Learning for Life*, particularly its stage of development and context, will influence the evaluation's purpose.

Four general purposes for an evaluation can be identified:

1. The purpose of an evaluation may be to gain insight. Knowledge from such an evaluation provides information concerning the practicality of a new approach, which can be used to design a program that will be tested for its effectiveness.

   This will not be the main purpose of the evaluation of *Learning for Life* as the main concern is not to investigate a new approach. The evaluation may, however, examine parts of the *Learning for Life* program with this purpose. These parts will need to be identified by stakeholders.

2. The purpose of an evaluation may be to change practice. Knowledge from this type of evaluation can be used to better describe program processes, to improve how the program operates, and to fine-tune the overall program strategy. Evaluations undertaken for this purpose include efforts to improve the quality, effectiveness, or efficiency of program activities. This is most appropriate in the implementation stage when an established program seeks to describe what it has done and to what extent.

   This may be a key purpose of the evaluation of *Learning for Life*.

3. The purpose of an evaluation may be to assess effects. Evaluations with this purpose examine the relationship between program activities and observed consequences.

   This type of evaluation is appropriate for mature programs that can define what interventions were delivered to what proportion of the target population.

   Knowing where to find potential effects can ensure that significant consequences are not overlooked. One set of effects might arise from a direct cause-and-effect relationship to the program. Where these exist, evidence can be found to attribute the effects exclusively to the program. In addition, effects might arise from a causal process involving issues of contribution as well as attribution. For example, if a program's activities are aligned with those of other programs operating in the same setting, certain effects cannot be attributed solely to one program or another. In such situations, the goal for evaluation is to gather credible evidence that describes each program's contribution in the combined change effort. Establishing accountability for program results is predicated on an ability to conduct evaluations that assess both of these kinds of effects.

   Assessment of effects may also be a key purpose of the evaluation of *Learning for Life* as it is a maturing program that can define its interventions.

4. The purpose of an evaluation may be to affect those who participate in the inquiry. This purpose applies at any stage of program development. The logic and systematic reflection required of stakeholders who participate in an evaluation can be a catalyst for self-directed change. An evaluation can be initiated with the intent of generating a positive influence on stakeholders. Such influences might be to:

   - supplement the program intervention (for example, using

	a follow-up questionnaire to reinforce program messages) • empower program participants (for example, increasing a family's sense of control over program direction) • promote staff development (for example, teaching staff how to collect, analyse and interpret evidence) • contribute to and facilitate organisational growth and development (for example, clarifying how the program relates to the organisation's mission) • facilitate social inclusion. This may be a subsidiary purpose of the evaluation of *Learning for Life,* given it will be conducted collaboratively with staff.
36.Users – the persons who will receive the evaluation findings.	Users are the persons who receive the evaluation findings. Because intended users directly experience the consequences of (inevitable) design compromises, they should participate in choosing the focus of the evaluation of *Learning for Life*. User involvement is required for clarifying intended uses, prioritising questions and methods, and keeping the evaluation on course.

37. Uses - the ways in which information from the evaluation will be applied.	Uses are the specific ways in which information generated from the evaluation will be applied. Stating uses in vague terms that appeal to many stakeholders increases the chances the evaluation will not adequately address anyone's needs. The uses of the evaluation of *Learning for Life* should be planned and prioritised with input from stakeholders and with regard for its stage of development and current context. All uses should be linked to one or more specific users.
38. Questions establish boundaries for the evaluation.	Questions establish boundaries for the evaluation by clarifying those aspects of the program to be addressed. Creating evaluation questions encourages stakeholders to reveal what they believe the evaluation should offer. Negotiating and prioritising questions among stakeholders further refines a viable focus. The question-development phase may also expose differing stakeholder opinions regarding the best units of analysis. For example, certain stakeholders might want to study how the *Learning for Life* program operates as an intervention to effect change within a particular community. Other stakeholders might have questions concerning the performance of a single element or a local group within it. Clear decisions regarding the questions and corresponding units of analysis are required to guide method selection and evidence gathering in subsequent steps of the evaluation.

39. Methods for gathering data will vary according to the questions being addressed.	Evaluation methods should be selected to provide the appropriate information to address stakeholders' questions; that is, methods should be matched to the primary users, uses and questions. The choice of design has implications for what will count as evidence, how that evidence is gathered, and what kind of conclusions can be made (including the internal and external validity[9] of findings). Decisions about methodology help to clarify the following: 1. to what extent program participants will be involved 2. how information sources will be selected 3. what data collection instruments will be used 4. who will collect the data 5. what data management systems will be needed 6. the appropriate methods of analysis, synthesis, interpretation and presentation. Because each method has its own limitations, evaluations that use more than one method are generally more effective. During the course of the evaluation, methods may need to be revised.
40. Agreement between stakeholders.	Agreement between stakeholders should: 1. summarise the evaluation procedures and clarify roles and responsibilities among those who will execute the evaluation plan and with those who authorise the evaluation 2. describe how the evaluation plan will be implemented with respect to resources (for example, money, personnel, time and information)

9 *Internal validity* refers to the extent to which the effects of the program are properly attributed to *Learning for Life*. *External validity* refers to the extent to which these findings generalise beyond *Learning for Life*.

3. state what safeguards are in place to protect human subjects and, where appropriate, the ethical or administrative approvals that have been obtained.

Agreements include statements concerning:

4. the intended purpose
5. users
6. uses
7. questions
8. methods.

Agreements include summaries of:

9. deliverables
10. timelines
11. budgets.

The agreement for the *Learning for Life* evaluation should be made between the primary users, any providers of financial or in-kind resources, and those persons who will conduct the evaluation and facilitate its use and dissemination.

The agreement for the *Learning for Life* evaluation should take the form of a detailed protocol. Creating such an explicit agreement verifies the mutual understanding needed for a successful evaluation. It also provides a basis for modifying or renegotiating procedures if necessary.

41. Standards for focusing the evaluation.	Each of the four standards give rise to questions which need to be addressed when focusing the evaluation of *Learning for Life*: 1. Utility a. What is the purpose of the evaluation? b. Who will use the evaluation results and how will they use them? c. What are special needs of any other stakeholders that must be addressed? 2. Feasibility a. What are the program's stages of development? b. How amenable to measurement are the components in the proposed focus? 3. Propriety a. Do the focus and design seek to adequately detect any unintended consequences of the program? b. Will the focus and design include examination of the experiences of those who are affected by the program? 4. Accuracy a. Is the focus broad enough to detect success, progress or failure of the program? b. Is the design able to respond to the questions – such as cause or attribution – that are asked by stakeholders? If not, why not?

42. Risks and risk management.	The following are seen as potential risks to Step 3: 8. <u>Risk</u>: in focusing the research design of the evaluation it may be difficult to ascertain whether *Learning for Life* is developing or maturing, if some elements of the program are at different stages. <u>Management of risk</u>: the evaluators will document reasons for decisions regarding the stage, or stages, of the program. These decisions should be made after consultation with stakeholders. Implications for the research design should form part of this consultation. 9. <u>Risk</u>: in focusing the research design and related methodology there appears to be a poor fit between the available resources and the best design. <u>Management of risk</u>: engage stakeholders, primary users and senior management in order to agree on the best fit between research design and available resources. This may involve adjustment to research design or changes in resourcing, or both. The evaluators will document any process of adjustment in the agreement for the evaluation. This will be accessible throughout the evaluation to help deal with questions concerning design and resourcing issues, and to keep the project on time and within budget.
43. Next steps.	On the following pages there are: • a checklist of activities that need to be undertaken to focus the evaluation of *Learning for Life* • worksheets to facilitate focusing the evaluation. The next chapter, *Step 4: Gathering credible evidence*, outlines approaches to the collection of data for the evaluation.

Checklist 3: Focusing the evaluation of Learning for Life

- [ ] Define the purpose(s) and user(s) of the evaluation.
- [ ] Identify the use(s) of the evaluation results.
- [ ] Consider stages of development, program intensity, logistics and resources.
- [ ] Determine the components of the logic model that should be part of the focus given the utility and feasibility considerations.
- [ ] Formulate the evaluation questions to be asked of the program components in the evaluation focus, that is, implementation, effectiveness, efficiency, and attribution questions.
- [ ] Review evaluation questions with stakeholders, *Learning for Life* managers, and *Learning for Life* staff.
- [ ] Review options for the evaluation design, making sure that the design suits the evaluation questions.

Worksheet 3: Focusing the evaluation in the logic model

If this is the situation …	**Then these are the parts of the logic model I would include in the evaluation focus**
1 Who is asking evaluation questions of *Learning for Life*?	
2 Who is expected to use the information from the evaluation questions and for what purpose?	
3 In Step 1 were the interests of other stakeholders that must be taken into account identified?	

Worksheet 4: Checking the evaluation focus

If this is the answer to these questions ...	**Then it is fair to conclude that the questions in the evaluation focus are or are not reasonable ones to ask now**
1 How long has *Learning for Life* been underway?	
2 How ambitious is *Learning for Life*? Is it multi-faceted or a simple intervention?	
3 What is the extent of the resources available for the evaluation?	

Step 4: Gathering credible evidence

The evaluation of *Learning for Life* will strive to collect information that conveys a well-rounded picture of the program so that the information and subsequent analyses are credible to the evaluation's primary users.

Credible evidence strengthens evaluation judgements and the recommendations that follow. Whether a body of evidence is credible to stakeholders depends on such factors as how the questions were posed, sources of information, conditions of data collection, reliability of measurement, validity of interpretations, quality control procedures and the receptiveness of stakeholders.

Encouraging participation by stakeholders enhances credibility. If stakeholders are involved in defining and gathering data that they find credible, they will be more likely to accept the evaluation's conclusions and act on its recommendations.

Although all types of data have limitations, the evaluation's overall credibility can be improved by using multiple procedures for gathering, analysing and interpreting data.

Aspects of evidence gathering that typically affect perceptions of credibility include indicators, sources, quality, quantity and logistics.

44. Indicators translate concepts to measures.	Indicators translate general concepts regarding the program, its context, and its expected effects into specific measures that can be interpreted. They therefore provide a basis for collecting evidence that is valid and reliable for the evaluation's intended uses.[10] Indicators address criteria that are used to judge the program and so reflect aspects of the program that are suitable for monitoring. Defining too many indicators can detract from the evaluation's goals; however, multiple indicators are needed for understanding the implementation and effects of *Learning for Life*. The approach to developing multiple indicators for the evaluation of

---

[10] *Validity* refers to the degree to which an instrument measures what it purports to measure of the information gathered. *Reliability* refers to the consistency of the information gathered using the same instrument.

Learning for Life should be based on the program logic model. The logic model should be used as a template to define a spectrum of indicators leading from program activities to expected effects. For each step in the model, qualitative or quantitative or both types of indicators should be developed to suit the concept in question, the information available, and the planned data uses.

Relating indicators to the logic model allows the detection of changes in performance more effectively than if a single outcome is the only measure used. Lines of responsibility and accountability are also clarified through this approach because the measures are aligned with each step of the program strategy. Further, this approach should result in a set of broad-based measures that reveal how outcomes are related to the direct and indirect effects of the program.

Measuring program performance by tracking indicators is only part of the evaluation and should not be seen as the sole basis for decision-making. Problems can result from using performance indicators as a substitute for completing the evaluation process and reaching fully justified conclusions. For example, an indicator might be assumed to reflect a failing or less than effective program when, in reality, the indicator might be influenced by changing conditions that are beyond the program's control.

The following observations assist in selecting indicators:

- indicators can be developed for activities (process indicators) or for outcomes (outcome indicators)
- there can be more than one indicator for each activity or outcome
- the indicator must be focused and must measure an important dimension of the activity or outcome
- the indicator must be clear and specific in terms of what it will measure.

45. The sources of evidence for the evaluation come from persons, documents and observations.	Sources of evidence in an evaluation are the persons, documents, records or observations that provide information for the inquiry. Where possible, more than one source should be used to gather evidence for each indicator to be measured. Selecting multiple sources provides an opportunity to include different perspectives regarding the program and thus enhances the evaluation's credibility. The criteria used for selecting sources should be stated clearly so that users and other stakeholders can interpret the evidence accurately and assess the extent to which it is appropriate. Sources for the evaluation of *Learning for Life* will be selected from or may include: • persons ○ *Learning for Life* families ○ staff, program managers, administrators ○ general public, sponsors, donors, corporate and community partners, the Board, the Research Advisory Committee ○ key informants ○ critics or sceptics ○ staff of other agencies ○ representatives of advocacy groups ○ elected officials, legislators and policymakers. • documents ○ grant proposals, newsletters, press releases, program documents, for example registration records, and reports drawn from the program ○ meeting minutes, administrative records, registration or enrolment forms

	○ publicity materials, quarterly reports, internet pages ○ publications, journal articles, posters ○ previous evaluation reports ○ asset and needs assessments ○ database records ○ records held by funding officials or collaborators. • observations ○ meetings, special events or activities, job performance.
46. Quality of the data refers to its reliability, validity and how informative it is.	Quality refers to the appropriateness and integrity of information used in the evaluation. Because all data have limitations, the intent of a practical evaluation is to strive for a level of quality that meets the stakeholders' thresholds for credibility. High-quality data are reliable, valid and informative for their intended use. Well-defined indicators enable easier collection of quality data. Other factors that affect quality include instrument design, data-collection procedures, training of data collectors, source selection, coding, data management, and routine error checking. Obtaining quality data entails balanced selections (for example, breadth versus depth) that should be negotiated among stakeholders in consultation with the evaluation team.

47. Quantity refers to the amount of data the evaluation needs.	Quantity refers to the amount of evidence gathered for the evaluation. The amount of information required should be estimated in advance, and where evolving processes are used, criteria should be set for deciding when to cease collecting data. Quantity affects the potential confidence level or precision of the evaluation's conclusions. It also partly determines whether the evaluation has sufficient power to detect effects. All evidence collected should have a clear, anticipated use. Correspondingly, only a minimal burden should be placed on respondents for providing information.
48. Logistics refers to the how and when of data collection.	Evaluation logistics encompass the methods, timing and physical infrastructure for gathering and handling evidence. Each technique selected for gathering evidence should be suited to the source(s), analysis plan, and strategy for communicating findings. Persons and organisations have cultural preferences that dictate acceptable ways of asking questions and collecting information. These include perceptions of who are appropriate persons to ask the questions. The procedures for gathering evidence in the evaluation should be aligned with the cultural conditions in each setting of *Learning for Life* and scrutinised to ensure that the privacy and confidentiality of the information and sources are strictly protected. In outlining procedures for collecting the evaluation data, the following issues need to be considered: • when to collect the data: The evaluation team should determine when (and at what intervals) it is most appropriate to collect information • who will be considered a participant in the evaluation • whether data will be collected from all participants or from a representative sample

	• who will collect the information, and related to this, how will the data collectors uniformly gather and record information. Data collectors will need to be trained to ensure that they all collect information in the same way and so reduce the risk of bias • how the security and confidentiality of the information will be maintained and related to this, what are the procedures for informed consent and data storage and retention.
49.Standards for gathering credible evidence.	Each of the four standards give rise to questions which need to be addressed when gathering credible evidence for the evaluation of *Learning for Life*: 1. Utility a. Have key stakeholders who can assist with access to participants been consulted? b. Are methods and sources appropriate to the intended purpose and use of the data? c. Have key stakeholders been consulted to ensure there are no untoward preferences for, or obstacles to, selected methods or sources? d. Are there specific methods or sources that will enhance the credibility of the data with key users and stakeholders? 2. Feasibility a. Can the data methods and sources be implemented within the time and budget for the project? b. Does the evaluation team have the expertise to implement the chosen methods? c. Are the methods and sources consistent with the culture and characteristics of the participants, such as language

and literacy levels?

   d. Are logistics and protocols realistic given the time and resources available for data collection?

3. Propriety

   a. Will data collection be unduly disruptive?

   b. How will issues of safety of respondents or confidentiality be addressed?

   c. Are the methods and sources appropriate to the culture and characteristics of the respondents – will they understand what they are being asked, and will they feel comfortable giving fully honest answers?

4. Accuracy

   a. Are appropriate quality assurance procedures in place to ensure quality of data collection?

   b. Are enough data being collected to support chosen confidence levels or statistical power?

   c. Are methods and sources consistent with the nature of the problem(s), the sensitivity of the issue(s), and the knowledge level of the respondents?

50. Risks and risk management.	The following is seen as a potential risk to Step 4: 10. Risk: in gathering evidence for the evaluation it becomes apparent that some desirable data are not extant, or not forthcoming. Management of risk: the evaluators will record the actions taken to secure or uncover data; The Smith Family senior management will support actions in data recovery; the evaluators will explain absences in the data utilised for the evaluation and how the evaluation managed these absences.
51. Next steps.	On the following pages there are: • a checklist of activities that need to be undertaken for credible evidence to be gathered for the evaluation of *Learning for Life* • worksheets to facilitate gathering credible evidence. The next chapter, *Step 5: Justifying conclusions*, outlines approaches to drawing conclusions, which are evidenced-based and rest upon agreed values and standards.

Checklist 4: Gathering credible evidence for the evaluation of Learning for Life

- [ ] Identify indicators for activities and outcomes in the evaluation focus.
- [ ] Determine whether existing indicators will suffice or whether new ones must be developed.
- [ ] Consider the range of data sources and choose the most appropriate.
- [ ] Consider the range of data collection methods and choose those best suited to the evaluation context and content.
- [ ] Pilot or test new instruments to identify and control sources of error.
- [ ] Consider a mixed-method approach to data collection.
- [ ] Consider quality and quantity issues in data collection.
- [ ] Develop a detailed protocol for data collection.

Worksheet 5: Evaluation questions, indicators and data collection methods and sources

Logic model components in evaluation focus	Indicator(s) or evaluation questions	Data methods and sources
1		
2		
3		
4		
5		
6		
7		
8		
9		

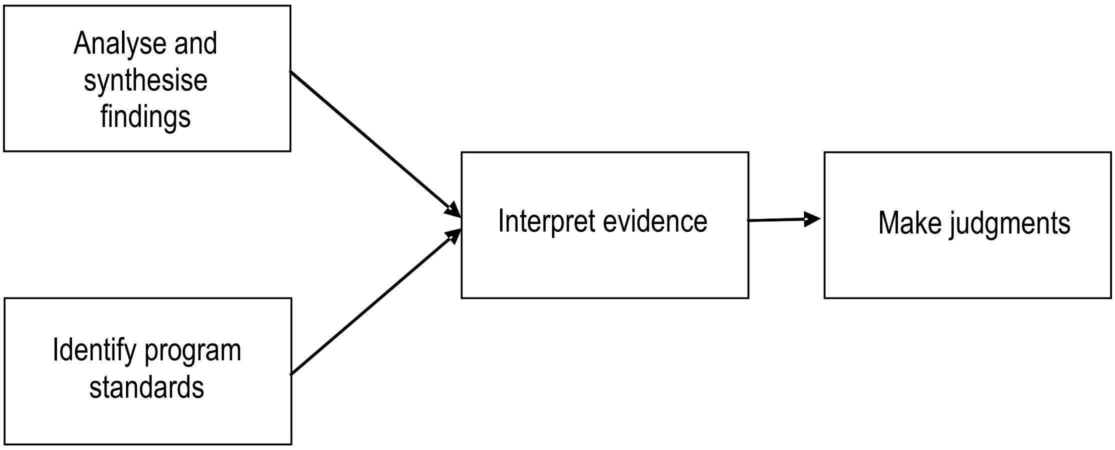
Worksheet 6: Data collection logistics

Data collection methods and sources	From whom data collected	By whom, and when will data be collected	Confidentiality and security steps
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			

Step 5: Justifying conclusions

Justification of conclusions is fundamental to sound evaluation.

The evaluation conclusions are justified when they are linked to the evidence gathered and judged against agreed-upon standards set by the stakeholders. Conclusions become justified when analysed and synthesised findings (the evidence) are interpreted through the values (standards) agreed upon by stakeholders and then judged accordingly. This process is summarised in the following diagram.



Stakeholders must agree that conclusions are justified before they will use the results of the evaluation with confidence. When stakeholders agree that the conclusions are justified, they will be more inclined to use the evaluation results for program development. Justifying conclusions on the basis of evidence for the evaluation of *Learning for Life* includes considering standards, analysis and synthesis, interpretation, judgment, and recommendations. These are understood in the ways described below.

52. Standards are used for judging performance of *Learning for Life*.	Standards reflect the values held by stakeholders and those values provide the basis for forming judgments concerning the performance of *Learning for Life*. In practice, when stakeholders articulate, negotiate and prioritise their values, these become the standards for judging whether a given program's performance will, for example, be considered successful, adequate or unsuccessful.

53. Analysis and synthesis are guided by questions asked, data available and input from key players in the evaluation.	Analysis and synthesis of an evaluation's findings aim to detect patterns in evidence, either by isolating important findings (analysis) or by combining sources of information and findings to reach a larger understanding (synthesis). Mixed method evaluations require the separate analysis of each evidence element and a synthesis of all sources for examining patterns of agreement, convergence or complexity. Deciphering facts from a body of evidence involves deciding how to organise, classify, interrelate, compare and display information. These decisions are guided by the questions being asked, the types of data available, and by input from stakeholders and primary users.
54. Interpretation is the attempt to understand the evidence.	Interpretation is the interrogation and summary of what the findings mean. It involves considering such questions as *How much is enough to be indicative of success? On what dimensions does change need to have occurred?* Interpretation is part of the overall effort to understand the evidence gathered in an evaluation. Uncovering facts regarding a program's performance is insufficient to draw evaluative conclusions. Evaluation evidence must be interpreted to determine the practical significance of what has been learned. Interpretations draw on information and perspectives that stakeholders bring to the evaluation inquiry and can be strengthened through active stakeholder participation. Appendix C lists some considerations and questions when interpreting evidence.
55. Judgments are statements concerning the merit, worth or significance of the program.	Judgments are statements concerning the merit, worth or significance of the program. They are formed by comparing the findings and interpretations regarding the program against one or more selected standards. Because multiple standards can be applied to a given program, stakeholders can reach different, or even conflicting, judgments. In the context of an evaluation, such disagreement can be a catalyst for clarifying relevant values and for negotiating the appropriate bases and standards on which the program should be judged.

56. Recommend-ations are action statements that flow from judgements about the program.	Recommendations are actions for consideration resulting from the evaluation. Forming recommendations is a distinct element of program evaluation that requires information beyond that necessary to form judgments regarding program performance. Recommendations for continuing, expanding, redesigning or terminating a program are separate from, but related to, judgments regarding a program's effectiveness. Making recommendations requires information concerning the context, particularly the organisational context, in which decisions will be made. Recommendations that lack sufficient evidence or those that are not aligned with stakeholders' values can undermine an evaluation's credibility. By contrast, an evaluation can be strengthened by recommendations that anticipate the political sensitivities of intended users and highlight areas that users can control or influence. Sharing draft recommendations, soliciting reactions from multiple stakeholders, and presenting options instead of directive advice increases the likelihood that recommendations will be relevant and well-received. Recommendations can be strengthened by: 1. summarising the plausible mechanisms of change 2. delineating the temporal sequence between activities and effects 3. searching for alternative explanations and showing why they are less well supported by the evidence. When different but equally well-supported conclusions exist, each may be presented with a summary of its strengths and weaknesses. Techniques for analysing, synthesising and interpreting findings should be agreed on before data collection begins to ensure that evidence gathered is able to be used.

57. Standards for justifying conclusions.

Each of the four standards give rise to questions which need to be addressed when justifying conclusions for the evaluation of *Learning for Life*:

1. Utility
    a. Have the perspectives, procedures and rationale used to interpret the findings been comprehensively described?
    b. Have stakeholders considered different approaches for interpreting the findings?
2. Feasibility
    a. Is the approach to analysis and interpretation appropriate to the levels of expertise and resources?
3. Propriety
    a. Have the standards and values of those less powerful, or those most affected by the program, been taken into account in determining standards for success?
4. Accuracy
    a. Can the conclusions be explicitly justified?
    b. Are the conclusions understandable to stakeholders?

58. Risks and risk management.	The following is seen as a potential risk to Step 5: 11. Risk: stakeholders do not accept conclusions and seek to stymie or censor these. Management of risk: the stakeholders are engaged throughout the evaluation and because of this necessary and effective communication, the evaluation's findings and conclusions should not be overly surprising. Senior management is supportive of findings and their widespread use. Processes for release and communications of findings and conclusions are agreed upon by stakeholders *before* the findings and conclusions have been noted.
59. Next steps.	On the following pages there are: • a checklist of activities that need to be undertaken for justifying conclusions arising from the evaluation of *Learning for Life* • a worksheet to facilitate justifying conclusions. The next chapter, *Step 6: Ensuring use of evaluation findings*, describes the final steps to be taken in the evaluation.

Checklist 5: Justifying conclusions in the evaluation of Learning for Life

- [ ] Check data for errors.
- [ ] Analyse data using appropriate techniques.
- [ ] Consider issues of context when interpreting data.
- [ ] Assess results against available literature and results of similar programs.
- [ ] If multiple methods have been employed, compare different methods for consistency in findings.
- [ ] Consider alternative explanations.
- [ ] Use existing standards as a starting point for comparisons.
- [ ] Compare program outcomes with those of previous versions (if possible).
- [ ] Compare actual with intended outcomes.
- [ ] Document potential biases.
- [ ] Examine the limitations of the evaluation.

Worksheet 7: Evaluation questions, indicators and data collection methods and sources

Question	Response
1 Who will analyse the data (and who will coordinate this effort)?	
2 How will data be analysed and displayed?	
3 Against what standards will interpretations be considered in forming judgments?	
4 Who will be involved in making interpretations and judgments and what process will be employed?	
5 How will conflicting interpretations and judgments be dealt with?	
6 Are the results similar to those expected? If not, why not?	
7 Are there alternative explanations for the results?	
8 If possible, consider how results compare with those of similar programs?	
9 What are the limitations of the data analysis and interpretation processes (for example, potential biases, generalisability of results, reliability, validity)?	
10 If multiple indicators were used to answer the same evaluation question, were similar findings obtained?	
11 Is it likely stakeholders will interpret the findings in an appropriate manner?	

Step 6: Ensuring use and sharing lessons

Lessons learned in the course of the evaluation of *Learning for Life* will not automatically translate into informed decision-making and appropriate action. Deliberate effort is needed to ensure that the evaluation findings are used and disseminated appropriately.

Five elements are critical for ensuring optimal use of the evaluation: design, preparation, feedback, follow-up and dissemination. These should be understood in the following ways.

60. Good design is fundamental to ensuring use.	Design refers to how the evaluation's questions, methods and overall processes are constructed. The design should be organised from the beginning so as to achieve intended uses by primary users. Having a clear design that is focused on use helps persons who conduct the evaluation to know precisely who should do what with the findings. The process of creating a clear design highlights ways that stakeholders, through their contributions, can enhance the relevance, credibility and overall utility of the evaluation.
61. Preparation is needed.	Preparation refers to the steps taken to get ready to eventually use the evaluation findings. Through preparation, stakeholders should: a. strengthen their ability to translate new knowledge into appropriate action b. discuss how potential findings might affect decision-making c. explore positive and negative implications of potential results and identify different options for program improvement and development.

62. Feedback helps to create trust.	Feedback is the communication that occurs among all parties to the evaluation. Giving and receiving feedback helps create an atmosphere of trust among stakeholders; it helps keep the evaluation on track by letting those involved stay informed about the progress of the evaluation. Primary users and other stakeholders have a right to comment on decisions that might affect the likelihood of obtaining useful information. Obtaining feedback will be encouraged by holding periodic discussions and utilising other communication forms as appropriate during each step of the evaluation process. Interim findings, provisional interpretations and draft reports will be shared routinely.
63. Follow-up is essential.	Follow-up refers to the technical and other support that users need during the evaluation and after they receive evaluation findings. Follow-up might be required to prevent lessons from becoming lost, or ignored, in the process of making complex or politically sensitive decisions. To guard against this a key person involved in the evaluation should be selected to serve as an advocate for the evaluation's findings during the decision-making phase. This type of advocacy should help increase appreciation of what was discovered and of actions that are consistent with the findings. Facilitating the use of evaluation findings also embodies the responsibility for working to prevent possible misuse. Evaluation results are bound by the context in which the evaluation was conducted. However, certain stakeholders might be tempted to take results out of context or to use them for purposes other than those agreed on. Active follow-up might help prevent misuse by ensuring that evidence is not misinterpreted or applied to areas other than those that were the central focus of the evaluation.

64. Dissemination should focus on relevant audiences and be full and impartial.	Dissemination is the process of communicating either the procedures or the lessons learned from an evaluation to relevant audiences in a timely, unbiased and consistent fashion. Regardless of how communications are structured, the goal for dissemination is to achieve full disclosure and impartial reporting. As with other elements of the evaluation, the reporting strategy should be discussed in advance with intended users and other stakeholders. Such consultation helps ensure that the information needs of relevant audiences will be addressed. Planning effective communication also requires considering the timing, style, tone, message source, vehicle and format of information products. A checklist of items to consider when developing evaluation reports forms Appendix D.
65. Standards for ensuring use of the evaluation.	Each of the four standards give rise to questions which need to be addressed when ensuring the use of the evaluation: 1. Utility a. Have significant interim findings and reports been shared with users so that findings can be used in a timely fashion? b. Has the evaluation been planned, conducted and reported in ways that encourage follow-through by stakeholders? 2. Feasibility a. Do reports clearly describe the program, including its context, and the evaluation's purposes, procedures and findings? b. Is the format of the reports appropriate to the evaluation's resources and to the time and resources of the audience?

	3. Propriety a. Have the evaluation findings (including the limitations) been made accessible to everyone significantly affected by the evaluation, and others who have the right to receive the results? b. Has all care been taken with respect to privacy and confidentiality in all the reporting processes? 4. Accuracy a. Have all reasonable attempts been made to avoid the distortions that can be caused by personal feelings and other biases? b. Do the evaluation reports impartially and fairly reflect evaluation findings?
66. Risks and risk management.	The following are seen as potential risks to Step 6: 12. Risk: stakeholders choose not to use the evaluation findings. Management of risk: the evaluators and The Smith Family management will engage and prepare stakeholders throughout the evaluation so there is trust in the strength of the evaluation and its findings, and therefore use is more likely. 13. Risk: stakeholders or others seek to modify, censor or distort evaluation findings. Management of risk: processes for communication and use of evaluation findings are agreed upon early in the evaluation process, and evaluators ensure that stakeholders are aware of this throughout the evaluation and beyond.

67. Checklists and worksheets follow.	On the following pages there are: • a checklist of activities that need to be undertaken for ensuring that the findings from the evaluation of *Learning for Life* are used and lessons are shared • a worksheet to facilitate the use and sharing of lessons learned.

Checklist 6: Ensuring that the findings from the evaluation of Learning for Life are utilised

- [ ] Identify strategies to increase the likelihood that evaluation findings will be used.
- [ ] Identify strategies to reduce the likelihood that information derived from the evaluation will be misinterpreted.
- [ ] Provide feedback to *Learning for Life* stakeholders.
- [ ] Prepare stakeholders for the eventual use of evaluation findings.
- [ ] Identify training and technical assistance needs.
- [ ] Use evaluation findings to support annual and long-range planning.
- [ ] Use evaluation findings to promote *Learning for Life* both internally and externally.
- [ ] Use evaluation findings to enhance the public image of the program and the organisation.
- [ ] Schedule follow-up meetings to facilitate the transfer of evaluation conclusions.
- [ ] Disseminate procedures and lessons to stakeholders.
- [ ] Provide interim reports to key audiences.
- [ ] Tailor evaluation reports to audience(s).
- [ ] Revisit the purpose(s) of the evaluation when preparing recommendations.
- [ ] Present clear and succinct findings in a timely manner.
- [ ] Avoid jargon when preparing or presenting information to stakeholders.

Worksheet 8: Communicating results

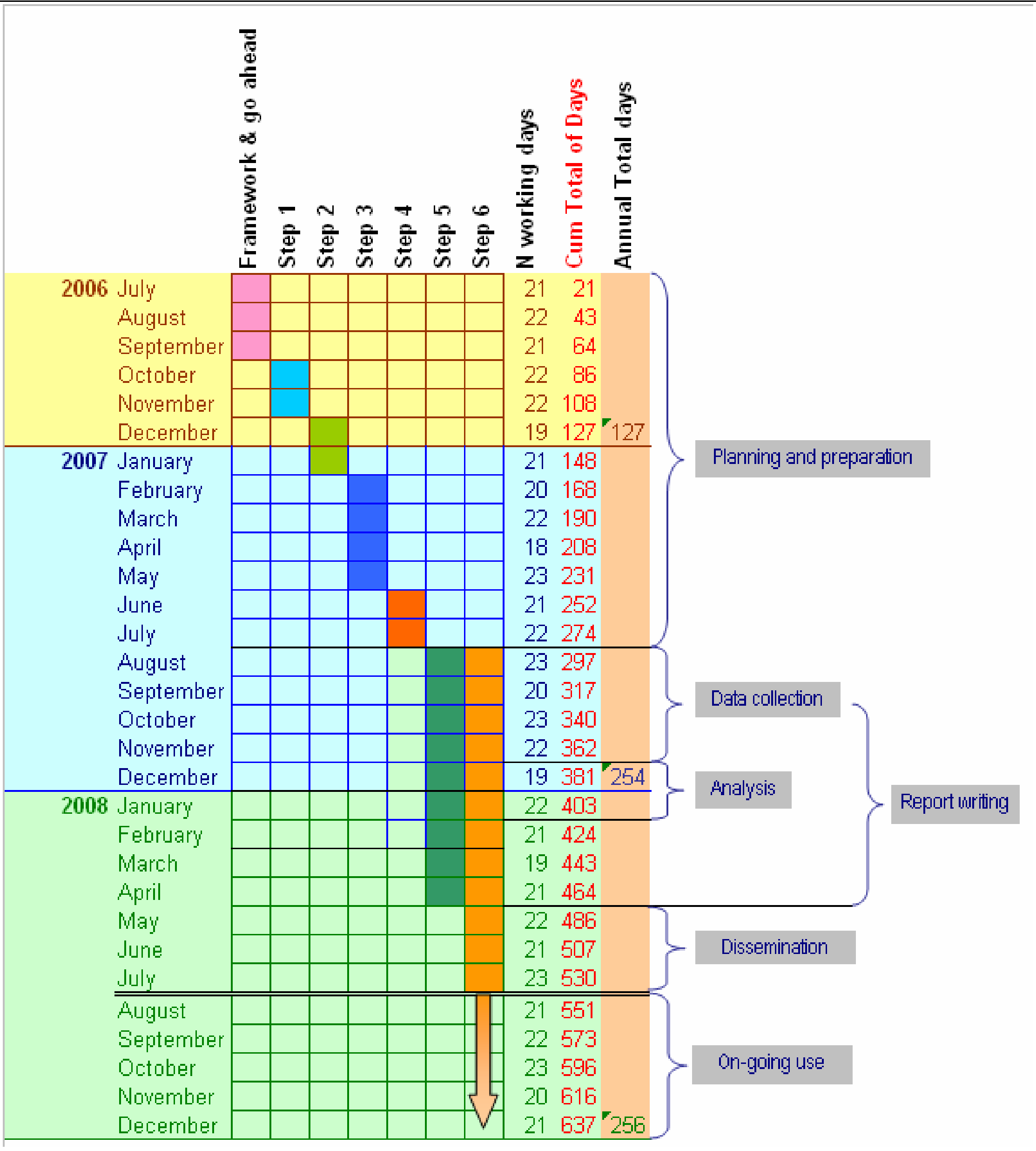
Need to communicate to this audience	The most appropriate format	The most effective channels
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Worksheet 9: Ensuring follow-up

Who will follow-up users of the evaluation findings?	How will follow-up be undertaken?	Support available for follow-up
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Timeline

Figure 1 shows a timeline for the conduct of the evaluation of *Learning for Life* broken down by the main steps to be taken. This figure is intended to provide a rough guide to the time required. Embedded in it are allowances for time for stakeholders to respond to requests for input and information.



Step 1 Engaging stakeholders
Step 2 Describing *Learning for Life* – development of a logic model
Step 3 Focusing the evaluation design – who, what, how and timelines
Step 4 Gathering credible evidence – instrument development
Step 5 Justifying conclusions – data analysis and interpretation, reporting
Step 6 Ensuring use and sharing lessons – reporting and dissemination

Figure 1 Summary timeline for the conduct of the evaluation of *Learning for Life*

Timeline

The key feature of the timeline is the thirteen-month planning and preparation period. This reflects the importance of meaningfully engaging stakeholders in the evaluation.

Appendix A: Selecting a lead evaluator and establishing an evaluation team

This appendix is designed to assist in selecting a lead evaluator and establishing an evaluation team.

a) An evaluation team is needed.	Good evaluation requires a combination of skills that are rarely found in a single person. An evaluation team that includes internal program staff, external stakeholders, and possibly consultants or contractors with evaluation expertise is the preferred approach. This should offer a balance of expertise and program and organisational knowledge.
b) The initial step – appoint a lead evaluator.	An initial step in the formation of a team is to decide who will be responsible for planning and implementing evaluation activities. At least one program staff person should be selected as the lead evaluator to coordinate program efforts. This person should be responsible for organising and overseeing evaluation activities, including: 1. planning and budgeting for the evaluation 2. developing objectives following appropriate consultation 3. addressing data collection needs 4. reporting findings 5. engaging stakeholders, consultants and other collaborators 6. working with consultants and others.
c) The Smith Family will have enhanced in-house evaluation capacity.	Developing and enhancing in-house evaluation expertise and capacity is a beneficial goal for The Smith Family.

d) Characteristics sought in a lead evaluator.

The lead evaluator should be able to work with a diverse group of stakeholders. This involves being willing and able to draw out and reconcile differences in values and standards of different stakeholders and to work with stakeholders in designing and conducting the evaluation.

Ideally, the lead evaluator also should:

1. have experience in the type of evaluation needed
2. be comfortable with qualitative and quantitative data sources and analysis
3. be able to work with a wide variety of stakeholders, including representatives of target populations
4. be able to develop innovative approaches to evaluation while considering the specificities and realities affecting the program (for example, a small budget)
5. understand both the potential benefits and risks of evaluation
6. educate program personnel about designing and conducting the evaluation
7. provide staff with comprehensive findings
8. have strong coordination and organisation skills
9. explain material clearly and patiently
10. respect all levels of personnel
11. communicate well with all personnel concerned
12. exhibit cultural competence
13. deliver reports and protocols on time
14. identify risks and changing influences and manage these appropriately.

e) External consultants.	External consultants can provide high levels of evaluation expertise from a more distanced and objective point of view. Important factors to consider when selecting consultants are: 1. their level of professional training 2. experience 3. ability to meet the needs of the evaluation 4. understanding, or capacity to understand the organisation and the sector where it operates. These should fit the program's evaluation needs and goals.
f) The evaluation team must reach consensus on key areas.	The evaluation team members should clearly define their respective roles. Informal consensus may be enough but a written agreement that describes who will conduct the evaluation and assigns specific roles and responsibilities to individual team members may be required. Either way, the team must clarify and reach consensus on the: 1. purpose of the evaluation 2. potential users of the evaluation findings and plans for dissemination 3. evaluation approach 4. resources available 5. protection for human subjects 6. timeline 7. budget 8. likely risks and risk management strategies.

Appendix B: Developing a logic model

Once the components of the program description have been identified, a visual depiction is a helpful way to summarise the relationship among any, or all, of the components. Logic models are graphic depictions of the relationship between a program's activities and its intended outcomes. Two words in this definition need emphasising: *relationship* and *intended*.

1. Relationship: Logic models convey not only the activities that comprise the program and the inter-relationship of those activities, but the link between these components and outcomes.

2. Intended: Logic models depict intended outcomes of a program's activities, rather than the reality at any point in time. As the starting point for evaluation and planning, the model serves as an outcomes 'roadmap' that shows the underlying logic behind the program, that is, why it should work. Of all activities that could have been undertaken to address this problem, these activities are chosen because, if implemented as intended, they should lead to the outcomes depicted.

The logic model requires no new thinking about the program, rather it converts the raw material generated in the program description into a picture of the program. This appendix provides steps for constructing and elaborating a simple logic model.

Constructing a simple logic model

Logic models may depict all or only some of the following components of the description of *Learning for Life*, depending on their intended use:

- inputs: resources that go into *Learning for Life* and on which it is dependent to mount its activities
- activities: events or actions undertaken in the conduct of *Learning for Life*
- outputs: direct products of *Learning for Life* activities, often measured in countable terms
- outcomes: the changes that result from *Learning for Life*'s activities and outputs, often in a sequence expressed as short-, medium-, and long-term outcomes.

A useful logic model can be constructed following these four steps.

A. Develop a list of activities and intended outcomes.

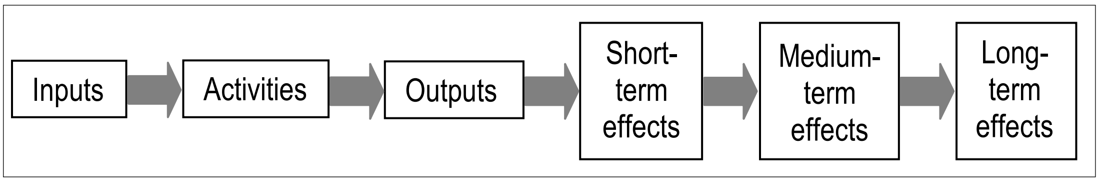
To stimulate the creation of a comprehensive list, any one of the following methods can be considered:

- Review any information available on *Learning for Life* – whether from its mission statement, strategic plans or key informants – and extract items that meet the definition of activity (something *Learning for Life* and its staff does) and of outcome (some change in someone or something, other than the program and its staff, that is hoped will result from the activities).
- Work backward from outcomes. This is 'reverse logic' logic modelling and may prove helpful when a program is given responsibility for a new or large problem or is just getting started. There may be clarity about the "big change" (most distal outcome) the program is to produce, but little else. Working backward from the distal outcome by asking "how to" helps identify the factors, variables and actors that will be involved in producing change.
- Work forward from activities. This is 'forward logic' logic modelling and is helpful when there is clarity about activities but not about why they are part of the program. Moving forward from activities to intended outcomes by asking, "So then what happens?" is often helpful in elaborating the downstream outcomes of the activities.

Logic models may depict all or only some of the elements of program description, depending on the use to which the model is being put. For example, Figure 2 is a simple, generic logic model. If relevant to the intended use, the model could include references to the remaining components of program description, such as context or stage of development. Adding inputs and explicit outputs to these examples would be a simple matter if needed.

Worksheet 10 is provided for recording the list of activities and outcomes.

Figure 2 Basic program logic model



B. Subdivide the lists to show the logical sequencing among activities and among outcomes.

Logic models provide clarity about the order in which activities and outcomes are expected to occur. To help provide that clarity, it is useful to take the single column of activities (or outcomes) developed in the last step, and then distribute them across two or more columns to show the logical sequencing. The logical sequencing may be the same as the time sequence, but not always. Rather, the logical sequence says, "Before this activity (or outcome) can occur, this other one has to be in place."

Worksheet 11 provides a simple format for expanding the initial two-column table.

C. Add any inputs and outputs.

The next step is to add columns for inputs and outputs. The inputs are inserted to the left of the activities while the outputs – as products of the activities – are inserted to the right of the activities but before the outcomes. This can be done using Worksheet 12.

Note that the outputs are the product of activities and should not be confused with outcomes.

D. Draw arrows to depict intended causal relationships.

The multi-column table of inputs, activities, outputs and outcomes may contain enough detail. However, when the model is used for planning and evaluation discussions, the logic model will benefit from adding arrows that show the causal relationships among activities and outcomes.

These arrows may depict a variety of relationships:

- from one activity to another, when the first activity exists mainly to feed later activities
- from an activity to an outcome, where the activity is intended to produce a change in someone or something (other than *Learning for Life*)

- from an early outcome to a later one, when the early outcome is necessary to achieve the more distal outcome.

Elaborating a simple logic model

A logic model is a picture depicting the program theory. It describes why *Learning for Life* should work.

Logic models can be broad or specific. The level of detail depends on the use to which the model is being put and the main audience for the model. A global model works best for stakeholders such as sponsors or corporate partners, but The Smith Family staff may need a more detailed model that reflects day-to-day activities and causal relationships.

When programs need both global and specific logic models, it is better to develop a global model first. The detailed models can then be seen as more specific magnification of parts of the program.

A simple logic model can be further elaborated in the following ways:

Elaborating distal outcomes

Sometimes the simple model will end with the short-term outcomes or even outputs. While this may reflect a program's mission, usually the program has been created to contribute to some larger purpose, and depicting this in the model may lead to more productive strategic planning discussions later. This elaboration is accomplished by asking "So then what happens?" of the last outcome depicted in the simple model, and then continuing to ask that of all subsequent outcomes until more distal ones are included. Elaborating the model in this way may lead to thinking more ambitiously about intended outcomes and whether the correct or appropriate activities are in place to produce these.

Elaborating intermediate outcomes

Sometimes the initial model presents the program's activities and its most distal outcome in detail, but with scant information on *how* the activities are to produce the outcomes. In this case, the goal of elaboration is to better depict the program logic that links activities to the distal outcomes. Providing such a step-by-step roadmap to a distal destination may help with:

1. identifying gaps in program logic that might not otherwise be apparent

2. persuading sceptics that progress is being made in the right direction, even if the destination has not yet been reached
3. aiding The Smith Family and *Learning for Life* managers in identifying what needs to be emphasised now and what can be done to accelerate progress.

The process of elaboration leads to the more detailed depiction of how the same activities *produce* the major distal outcome, that is, the milestones along the way.

Worksheet 10: Raw material for logic model

Activities	What changes will result in someone or something (other than *Learning for Life* or its staff)?
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Worksheet 11: Sequencing activities and outcomes for the logic model (1)

Activities		Outcomes	
Early	Later	Early	Later
1			
2			
3			
4			
5			
6			
7			

Worksheet 12: Sequencing activities and outcomes for the logic model (2)

Activities			Outcomes		
Inputs	Early	Later	Outputs	Early	Later
1					
2					
3					
4					
5					
6					
7					

Appendix C: Points to consider when interpreting findings

Judgments are statements about a program's merit, worth or significance. They are formed when findings are compared against one or more selected program standards. The following is a list to consider when interpreting findings:

- Interpret evaluation results with the goals of the program in mind.
- Keep audiences in mind when preparing the report. What do they need and want to know?
- Consider the limitations of the evaluation:
    - possible biases
    - validity of results
    - reliability of results
- Are there alternative explanations for the results?
- How do the results compare with those of similar evaluations?
- Have the different data collection methods used to measure your progress shown commensurate results?
- Are the results consistent with theories supported by previous research?
- Are the results similar to those expected? If not, why not?

Appendix D: Checklist for ensuring effective evaluation reports

1. Provide interim and final reports to users in time for effective use.
2. Tailor the report content, format and style for the audience(s) by involving audience members.
3. Include report summaries.
4. Summarise the description of the stakeholders and how they were engaged.
5. Describe essential features of the program.
6. Explain the focus of the evaluation and its limitations.
7. Include an adequate summary of the evaluation plan and procedures.
8. Provide all necessary technical information.
9. Specify the standards and criteria for evaluative judgments.
10. Explain the evaluative judgments and how they are supported by the evidence.
11. List both strengths and weaknesses of the evaluation.
12. Discuss recommendations for action with their advantages, disadvantages and resource implications.
13. Ensure protections for students, families and other stakeholders.
14. Anticipate how people or organisations might be affected by the findings.
15. Present minority opinions or rejoinders where necessary.
16. Verify that the report is accurate and as unbiased as possible.
17. Organise the report logically and include appropriate details.
18. Remove technical jargon.
19. Use examples, illustrations, graphics and stories.
20. Be concise.

Appendix E: Planned products from the evaluation of Learning for Life

There are three types of products planned to be produced by the evaluation of *Learning for Life*:

1. internal working documents, which are documents of limited circulation for use by the evaluators and stakeholders in the conduct of the evaluation
2. a formal agreement between stakeholders, which is to function as a memorandum of understanding between them
3. reports, which are prepared for public use.

In total, nine main products are to be produced by the evaluation of *Learning for Life*. These include:

1. An internal working document titled: *A list of persons and organisations identified as stakeholders in the evaluation of* Learning for Life. This document will list all stakeholders and provide a rationale for their inclusion on this list.
2. An internal working document titled: *A description of the* Learning for Life *scholarship program.* This document will provide a description of the program and will include a completed logic model for *Learning for Life.* This document will explain the contents of the logic model and its implications for the evaluation. The final version of this document will be subsumed into the final report.
3. A formal document named: *Agreement concerning the evaluation of the* Learning for Life *scholarship program.* This document will contain the contents of the agreement between all stakeholders participating in the evaluation.
4. An internal working document titled: *The piloting of instruments for the evaluation of* Learning for Life. This document will include a description of the methodology, the items selected for use and the rationale for the item selection. Sections of this document will be subsumed into the technical report.
5. An internal working document titled: *Data collection instruments for the evaluation of* Learning for Life. This document will contain the full set of data collection instruments used in the evaluation plus a rationale for each item in each data collection instrument. Sections of this document will be subsumed into the technical report.

6. An internal working document titled: *Data collection protocols for the evaluation of the* Learning for Life *scholarship program* which will describe how data are to be collected, processed and stored. It will also include the code books used for classifying data. Sections of this document will be subsumed into the technical report.

7. A report titled: *Technical report for the evaluation of The Smith Family's* Learning for Life *scholarship program*. This document will contain the technical details of all aspects of the conduct of the evaluation including sampling, item development and characteristics, quality control and monitoring procedures and field operations.

8. A report titled: *A report on the evaluation of The Smith Family's* Learning for Life *scholarship program.* This document will provide a description of the evaluation with its audience being the key stakeholders.

9. A report titled: *Summary report of The Smith Family's* Learning for Life *scholarship program*. This document will provide a description of the evaluation with its audience being all stakeholders and other interested parties, including the commercial media.

Sundry interim reports are also planned which will describe progress at key times in the evaluation. Other forms of communication, which may be construed as products such as presentations and seminars, press releases and journal articles, may also be produced, especially during the final stages of the evaluation.