FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

Evaluating Injury Prevention Initiatives

RESEARCH CENTRE FOR INJURY STUDIES
FLINDERS UNIVERSITY
Evaluating Injury Prevention Initiatives:
An Annotated Bibliography

Developed by:
FLINDERS INSTITUTE OF PUBLIC POLICY AND MANAGEMENT
Latest Revision May 1999

FOR FURTHER INFORMATION AND COMMENT

This document is open for comment and suggestions for further annotated abstracts. For further information and comment:

Dr James Harrison, Director
AIHW National Injury Surveillance Unit, in the
Research Centre for Injury Studies
Flinders University Research Centre for Injury Studies
Mark Oliphant Building; Laffer Drive; Bedford Park SA 5042

ACKNOWLEDGMENTS

This project was supervised by former RCIS Assistant Director, Mr Jerry Moller in March, 1997, and RCIS Director, Associate Professor James Harrison from May, 1997. The FIPPM staff responsible were Associate Professor Colin Sharp and John Stanwick.

This annotated bibliography benefited from the patience and assistance of many professionals dedicated to evaluation in the various injury prevention fields.

We especially thank the following for their contribution:

Brian Adams (Accident Injury Prevention Unit, Wellington, New Zealand)
Andrew Dannenberg (Centres for Disease Control and Prevention, Atlanta Georgia);
Jane Elkington (Consultant, Sydney);
Les Fisher, Archivist, American Public Health Association (Injury Control and Environmental Health Services Section)
Lynne Fragar (Australian Agricultural Health Unit, Moree, New South Wales)
Richard Franklin (Australian Agricultural Health Unit, Moree, New South Wales)
Paul (Mike) Furbee, Co-manager of INJURY-L listserver, (Center For Rural Emergency Medicine, West Virginia University)
James Harrison (National Injury Surveillance Unit, Flinders University, Adelaide, South Australia)
Michael Hill (Queensland Health, Brisbane, Queensland)
John Langley (Injury Research Unit, University of Atago, Dunedin, New Zealand)
Merrian McCormick (National Injury Surveillance Unit, Flinders University)
Jerry Moller (Commonwealth Department of Health and Community Services, Adelaide, South Australia)
Jim Nixon (Child Health Unit, University of Queensland, Brisbane, Queensland)
Joan Ozanne-Smith (Accident Research Centre, Monash University, Melbourne, Victoria)
Pierre Phillippe, Co-manager of EPIDEMIO-L listserver (Preventative Medicine, University of Montreal)
Leon S. Robertson (Department of Epidemiology and Public Health, Yale University, New Haven, Conn.)
Yvonne Robitaille, (Department of Community Health, Montreal Canada)
Ellie Rosenfeld (Southern Community Health Research Unit, Adelaide, South Australia)
Christine Roserveare (Hutt Valley Health, New Zealand)
Gary Sorock (Liberty Mutual Research Centre for Safety & Health, Hopkinton, Massachusetts)
Hassan Soubhi, Co-manager of EPIDEMIO-L listserver, (University of Montreal)
Diane Thompson, (Harbourview Injury Prevention & Research Centre, University of Washington)
John Wall (Road & Traffic Authority, Wollongong, New South Wales)
Jerome Winston, Program for Public Sector Evaluation, RMIT University. Melbourne, Australia; Co-manager of GOVTEVAL listserver,
Craig Zwerling (University of Iowa Injury Prevention Research Centre, Iowa)

Our thanks also to those, not listed here, who contributed to this document.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

CONVENTIONS USED HERE

As a resource to those involved in the injury surveillance and prevention field, we have used both conventional Harvard (Authors, date) citation of references, as well as electronic internet and World Wide Web (URL) addresses, wherever possible.

Australian Sources

As this review has been conducted by Australians, we have been able to access both the international and the Australian literature. Unfortunately the latter does not often reach overseas readers. Consequently we have indulged in an intended bias to use Australian sources where ever possible.

The classifications schemes used might have been based on the Australian *Human Services Thesaurus* (Department of Human Services and Health, 1995); and the Australian Bureau of Statistics (ABS, 1996) *Research Classification Schemes.* However, unfortunately, they were found inadequate for the purpose. It is clear that there are some gaps[1] and inconsistencies[2] in these research codes as far as the foci of this bibliography are concerned. The only advantage in adopting such a standard research classification scheme would have been the comparability with the ABS and OECD databases. However, this did not warrant the effort in making the various models and types of evaluation and injury prevention fit into such a small undiscriminating set of categories. Again the ABS classification scheme highlights the difficulties in data collection processes in the field of injury prevention.

[1] For example, *injury prevention* and *evaluation methodologies* are not listed as categories

[2] Compare the classification of *"environmental health"* as a *Public health* research field (100801) along with "*environmental and occupational health* and safety" and in terms of a Division 3 *Public Health* Socio-economic objective classification as *130207 Environmental health* **and** under Division 4 "Social Environment" (code 160604).

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

INTRODUCTION

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

TOOLS FOR DESIGNS & CLASSIFICATIONS OF EVALUATIONS OF INJURY PREVENTION

At the point of designing an injury prevention program is the ideal moment to plan for the linkage between the program and the evaluation. The design of the evaluation should be integrated into the design of the program. It is this connection which strengthens the degree of control over the implementation and the interpretation of the consequences of a program. Leaving the preparation of the evaluation as an afterthought is like closing the door after the horse has bolted. The opportunity to collect useful data and to use it to understand or improve the forces working in the program or system will be limited or lost.

A significant component in the program logic approach to evaluation is the choice of methods and the linkage between the evaluation questions and the data collection and analysis.

Like epidemiology (e.g. Sackett et al, 1985), and injury prevention research more generally (e.g., Harrison & Cripps, 1994), evaluation often draws on logical and statistical inference in the search for causal linkages in the identification of which program or intervention factors may or may not have contributed to the outputs and outcomes observed. However, even one of the more regulated and higher resourced areas of the injury prevention, that focusing on occupational injury in the workplace (Zwerling et al. 1997), there may be difficulty conducting highly rigorous evaluation based on experimental research techniques. Zwerling et al. (1997, p. 164) suggest that in evaluation of occupational injury prevention randomised controlled studies are "rare". Indeed, reflecting on the state of practice and the practical impediments to ideal methods, they recommended:

> "a hierarchical approach to evaluating occupational injury interventions - beginning with qualitative studies, following up with simple quasi-experimental designs using historical controls, continuing with more elaborate quasi-experimental designs" (Zwerling et al 1997, p. 164)

No doubt there are significant problems in applying experimental designs in most health promotion (see Hollister, & Hill, 1995) and injury prevention fields (see Dannenberg et al, in press). However, as one editorial in the health promotion field pointed out:

> health promotion evaluation "does however, need to think about rigour and scientificity [sic] if it is to be taken seriously as a major contributor to health care outcome strategies and to the evaluative research needed to support it"
> (Macdonald, 1996, p. 171)

Similarly, one could argue the case for this view in the injury prevention evaluation field. The three points (three evaluative strands) McDonald recommended are worth repeating in the present context:

> "The **first** strand would acknowledge that quasi-experimental research alone cannot provide the answers to many of the questions that we would need to know ... It might be that the field of evaluative research would be better served by the judicious use of descriptive, perhaps qualitative

> research methods that involve process or formative studies. This would help us understand issues to do with strategic and programme development: resource allocation and distribution; and unanticipated consequences of interventions. This approach may give better information and results for policy-makers and fund-holders than imperfect (quasi) experimental studies.
> However, to do this as rigorously as possible it might be that we need to construct a qualitative research methods hierarchy which puts the best, most theory-generating, most valid method at the top (to equate with RCTs on the quantitative hierarchy) and the more descriptive and less inductive studies at the bottom.
> **Secondly,** we need to look more closely at developing intermediate and indirect indicators that provide us with assessment data not readily available through experimental-outcome-focused studies. ...
> **Thirdly,** we need to think about combining intermediate indicators of success with longer term outcomes effectiveness or, ... of combining good quantitative and qualitative evaluative research "
> (Macdonald, 1996, p. 172, emphases original)

Indeed in the evaluation literature there are plentiful examples and models to address these concerns, they obviously need to be better promoted and implemented in the health promotion as well as the injury prevention fields. For example, Patton (1987; 1990) has long advocated mixing and complementing qualitative and quantitative methods for the reasons Macdonald mentions, as well as for the increased validity which comes from triangulation (Patton, 1987; 1990) of alternative sources. This approach has not been missed in the health promotion in Australia. For example, the evaluation of the Healthy Cities Noarlunga Project (Baum, 1992; Cooke & Baum, 1991) has been able to integrate various quantitative methods into the largely qualitative orientation. This excellent longitudinal program has also produced some valuable resources which will be reviewed in the annotated bibliography sections below.

It is easy to agree with Macdonald's (1996) point about the need for some appreciation of the relative strength or interpretative power of qualitative methods, but in terms of the framework and assumptions of the qualitative approaches the hierarchy of evaluation design might be more inclined to be the reverse of the view from the quantitative perspective.

Although there are barriers to `conducting highly vigorous evaluation based on experimental research techniques', it is important not to underestimate the potential for (or importance of) *efficacy* studies and *effectiveness* studies. Part of the approach should be careful attention to distinguishing planning, resourcing and management questions from research and efficacy questions. The former may be quite amenable to low cost, less complicated, less rigorous quasi-experimental designs. They allow demonstrably inefficacious or *harmful* methods to be avoided, and can improve the relatively certain "components" of the program in which other components may be less well known or less interpretable. For example, large resource allocation to widespread prevention programs may hinge on questions concerning whether we can demonstrate that the proposed intervention/program is effective enough to warrant the expense. Also once such programs are initiated, close attention to integrating data gathering, monitoring and *formative* evaluation techniques may avoid wastage or unexpected side-effects of the program, thus making the most of the investment while keeping it on track.

The prior research and efficacy questions may require more exacting or artificial experimental or laboratory approaches, under constrained resourcing and piloting conditions, which have to be undertaken to demonstrate that the intervention not only does produce a desired effect, but also *does no harm*.

Thus the type of issue or question raised, the stage of development of the intervention, and the resourcing available for the implementation, are all factors to be considered in deciding how to plan and manage an evaluation in conjunction with the development of the intervention.

The present review will examine the options in such an "hierarchical approach" to the choice of methods and the design of data collection. The following sections introduce a rudimentary set of flow charts for advice on design and classification of evaluation foci. These foci are a significant key to the coding used for the annotated bibliography (see category B1 Code segments). The basic framework for these charts has been developed from the classical analysis by Rossi and Freeman (1989), and integrated with other useful approaches and design tools, such as the excellent needs analysis framework provided by McKillip (1987, pp 102 -103), and the USA Government's General Accounting Office's (GOA/PEMD-10.1.4 1991, pp. 68 - 69 http://www.gao.gov/policy/10_1_4.htm) excellent guide to evaluation design. Also important are the linkages with the planning of the intervetion, so Charts V and VI give an overview of the connections between the evaluation plan and epidemiology of the intervention and management of the program.

The entry point to these charts and an overview of the whole framework is through Figure 7. The other charts are optional and are all linked through the simple steps outlined in Figure 7, available here as a PDF file (approx 21 Kb).

Figure 7 provides an overview for Planners and Managers, who have become involved with the design and management or an injury prevention. How do you link this with planning of the evaluation, as recommended above? Beginning at the point of the decision to implement some intervention (which will be called X from now on) which has some aims or objectives (lets call them Y)[9]. The Charts can be used to followup the detail of the general plan outlined in Figure 7. Chart I (PDF, approx 23 Kb) (on linking the evaluation question to evaluation design) begins as early in the planning (Figure 7) as is necessary to address the issues raised about the intervention. Chart II (PDF, approx 20 Kb) steps back from Chart I and Figure 7 to the prior question of identifying a problem in the logic of the injury intervention. Chart III (PDF, approx 31 Kb) follows from Chart II to address planning and management issues. Chart IV (PDF,approx 43 Kb) gives the main detail on the choice of evaluation designs. Chart V (PDF, approx 18 Kb) steps back to the basic factors and models of injury prevention and how they relate to the planning and evaluation context. The next sections introduce these Charts.

---

[9] Other Charts start with differing assumptions. Chart VI (PDF, approx 32 Kb) is more appropriate if the decision has still not been made, or a proposal for an intervention has to be evaluated first.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

METHODOLOGY

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

GENERAL REFERENCES

In the reference lists which follow there are overlaps between contents of some of the lists.
The lists are as follows:

- references cited in the introduction;
- references on evaluation, theory and practice
- references on evaluation of injury surveillance and/or injury prevention
- WWW resources & internet (email) addresses

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

1.1 THE NEED FOR A BIBLIOGRAPHY ON EVALUATION OF INJURY PREVENTION

From time to time disasters or `human interest stories' bring the risk of injury and the associated injury prevention policies and consequences to the attention of various communication media, or Governments and community agencies, locally and internationally[3]. As a recent review of childhood injury prevention put it:

> "Worldwide, injury is a leading cause of death in childhood and a major cause of morbidity and long-term disability. The implementation of prevention strategies of proven efficacy is of major public health importance. Finding out "what works" in injury control is of tremendous public health importance. Systematic literature reviews and meta-analyses are invaluable methods of synthesizing the existing evidence from evaluation studies. Health care providers, policy makers and injury control professionals are faced with large amounts of information, distributed in a large number of sources, and need systematic reviews to provide a basis for rational decision making. It is quite likely that when currently available evidence on the efficacy of injury prevention interventions is thoroughly synthesized, many interventions believed to be effective will be shown to be ineffective and vice versa. In addition, systematic reviews are likely to show that some proposals for future research are redundant because intervention efficacy can already be established from existing evidence. Most importantly, such reviews will clarify which programs are appropriate to implement on a broader scale."
> (Rivara, Beahler, Patterson, Thompson, & Zavitkovsky, 1997; http://weber.u.washington.edu./~hiprc/childinjury/)

How can the professionals, and the decision-makers, or representatives of community and Government, improve their response to such need for information and decision? How can we learn to more effectively control the problem of injury?

One avenue of opportunity is to expand the funding for injury prevention research and preventative programs. Another option is the collection and dissemination of more accurate, reliable and useful information about the risk and incidence of injury, to provide a better basis for setting priorities, and monitoring change in injury occurrence.

But would this effort and expenditure achieve what it sets out to do? Unless there is a concerted effort to evaluate the effectiveness of the use of the resources allocated to injury prevention such questions will remain a matter of conjecture.

Accordingly, recent commentators on injury prevention, recommend that evaluation should be an integral part of any injury prevention strategy:

> "An injury intervention should be evaluated to show it prevents injuries in the target population, to identify unintended consequences, to correct any problems that limit effectiveness, to justify resources from funding agencies, to seek continuing resources from funding agencies, to guide replication of the intervention elsewhere, and to prevent wasted resources if not effective."
> (Dannenberg, 1996, p.1;)

No doubt this is a highly laudable recommendation, however, defining the scope of evaluation for an injury prevention project, and selecting appropriate methods, are not trivial tasks. The term "evaluation" can include many approaches to data collection and decision-making, ranging from simple, brief and inexpensive to complex, time-consuming and expensive. Depending on circumstances, any of a wide range of evaluation methods may be appropriate to injury prevention programs.

In recognition of the need for evaluation in the injury prevention field, the Flinders Institute of Public Policy and Management (FIPPM) was asked to develop an annotated bibliography according to a brief provided by the National Injury Surveillance Unit (NISU). The purpose was to lift the level of resources available in the evaluation of injury prevention. It was not intended to gather articles dealing exclusively with any injury prevention issue per se, but to find the relevant evaluation methods, theory and applications in related fields which would be applicable to the work of injury prevention. The brief expected that perhaps 10 key articles in each of several key aspects of evaluation of injury prevention would be included in the final annotated bibliography. The emphasis was on quality rather than quantity. There was also an expectation of diversity and innovation of approaches to be canvassed.

As the literature was reviewed, the need emerged to add to the breadth of the categories to encompass the diversity of the set of the intersection of injury prevention and program evaluation. Hence the number of papers reviewed expanded.

---

[3] For example, in the USA see the National Agenda for Injury Control, 1992; National Committee for Injury Prevention and Control 1989; in Australia (e.g., Child Health Program of Queensland Health; National Health Priority Committee, Victorian Health Promotion Council) and internationally (e.g., media coverage of the blood alcohol level of the driver in the car crash in Paris which killed Princess Diana; INJURY-L discussion in September, 1997 of the non-use of seat belts of those that died in that car).

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

1.2 FORMAT OF THE BIBLIOGRAPHY

This overview will begin with an introductory framework to the field of evaluation, before addressing the methodology of the bibliography. Then three bibliographies are provided:

1. the references cited in the introductory sections are listed alphabetically and categorised in terms of the content area ;
2. references used in the annotated bibliography are listed alphabetically by author and numbered in terms of the year and a general Ref # (see Part D);
3. the selected annotated bibliography is provided in various sections, (see Part E).

The aim of the introductory remarks is to provide a common frame of reference for the bibliography. There is insufficient scope here for a comprehensive review of evaluation. Indeed, there is no need to duplicate the many useful texts (e.g., Attkisson, Hargreaves, Horowitz, & Sorensen, 1978; Guttentag. & Struening, 1975; Hawe, Degeling & Hall, 1990; Love, 1994; Owen, 1993; Patton, 1997; Rossi & Freeman, 1993; Scriven, 1991; Wholey, Hatry, & Newcomer, 1994); or introductory handbooks cum `do-it-yourself' guides to evaluation (e.g., Abbott & Craig, 1994; Australian Youth Foundation & Sharp, 1996; Herman, Morris & Fitz-Gibbon, 1988; Isaac & Michael, 1995; Linney, & Wandersman, 1991; Kemmis, 1985; Wadsworth, 1984; 1991); or specialised technical guides in evaluation and social research methodologies (e.g., see Fink 1995). Each of these types of reference are separately categorized in the general reference bibliography for accessibility. For example, there are extensive literatures and substantial texts devoted to methodologies related to evaluation, such as *needs analysis* (e.g., Baum, 1992; McKillip, 1987; Percy-Smith, 1996; South Australian Community Health Research Unit, 1991; 1996), *experimental and quasi-experimental techniques* (e.g., Boruch, 1997; Campbell & Boruch, 1975; Campbell & Stanley, 1963; Cook & Campbell, 1973, Cronbach, 1982; Lewis-Beck, 1993) and *social survey techniques* (see Dillman, 1991; 9 volumes edited by Fink, 1995; Fowler, 1993, 1995; Moser & Kalton, 1981; South Australian Community Health Research Unit, 1994). The intention here is not to provide detailed instruction in such techniques, but to point the reader in the direction from whence such guidance can be obtained.

We have not attempted to canvas the associated fields of economic evaluation of health and health program evaluation broadly (see eg., Hall[4], 1995). There are quite specific skills and a distinctive literature which are more appropriately accessed elsewhere (e.g., see the Sydney University based Centre for Health Economics Research and Evaluation, and the Melbourne based Centre for Health Program Evaluation,[5] the U.K. NHS Economic Evaluation Database[6]).

Of particular relevance are the excellent *Australian* resources available. In the field of evaluation of community based programs, there have emerged at least three practical evaluation kits (see AYF & Sharp, 1996; South Australian Community Health Research Unit, 1994 - 1996; Wadsworth, 1984, 1991). Indeed, in a broader sense the text by Hawe, Degeling and Hall (1990) has largely paved the way for the present annotated bibliography; both in terms of an insightful theoretical review and in the form of a useful collection of abstracts of key papers and books, most of which are still pertinent today. They have covered many of the issues which will be referred to here, but from the more general perspective of health promotion. Here, apart from the need for an update in the currency of the literature, and access to internet resources, the present review is an attempt to deal with both the macro level of an introduction to evaluation in general, as well as the micro level in terms of evaluation of injury prevention, per se.

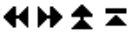
1.2.1 Intended Audience

The aims of this annotated bibliography are: to assist primarily those who are beginning to involve themselves in evaluation of injury prevention and control programs[7]; and to promote evaluation as a useful tool, complementary to injury prevention program management. The basis for this approach is the need for administrators and practitioners to be able to draw on a wider range of resources for evaluation, in order to facilitate the useability of evaluations and strengthen the credibility of recommendations to reduce injury. Unfortunately, there are relatively few main stream references to evaluation in the injury prevention and health promotion literature (see e.g., Harrison, & Cripps, 1994). We hope that the general sources in the injury prevention and health promotion literature will focus more on evaluation as some of those in these fields have recently recommended (see e.g., Danneberg, 1996; Miller, 1996; Towner, 1996; Zwerling, Daltroy, Fine, Johnston, Melius, &. Silverstein, 1997).

---

[4] http://www.health.usyd.edu.au/people/halljane.htm
[5] Department of Public Health and Community Medicine, University of Sydney, see http://www.health.usyd.edu.au/chere/
Department of Public Health and Community Medicine, University of Melbourne with the Faculty of Business & Economics at Monash University, see http://ariel.unimelb.edu.au/~chpe/
[6] The National Health Service Economic Evaluation Database uses telnet protocols, but can be accessed via their WWW site through http://www.york.ac.uk/inst/crd/info.htm
via University of York's NHS Centre for Reviews and Dissemination as part of their Cochrane Collaboration, see later Table 4.
[7] Our audience may not be newcomers to injury surveillance or injury prevention per se, nevertheless, we have provided accessable linkages into the injury prevention literature for those who are new to it; and also because of the need for an integrative approach to evaluation and program management.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit



1.3 APPROACH TO EVALUATION

Evaluation has been defined in many ways; we will examine three relevant definitions of the term, as follows:

1.3.1 Evaluation as a discipline of inquiry

Evaluation has been defined as

> "the process of determining the **merit**, **value** or **worth** of something, or the product of that process" (Scriven, 1991, p. 139)
> it is "the most powerful and versatile of the 'transdisciplines' - tools disciplines such as logic, design, and statistics - that apply across broad ranges of the human investigative and creative effort ..."
> (Scriven, 1991, p. 1)
> "the evaluation process normally involves some identification of relevant standards of merit, worth or value; some investigation of the performance of the evaluands on these standards; and some integration or synthesis of the results to achieve an overall evaluation ..." (Scriven, 1991, p. 139)

1.3.2 Program evaluation

The most common usage of evaluation is the context of a coordinated set of projects or services of interventions clustered around a coherent set of goals.

Program evaluation can be defined as follows:

a. A process of making reasonable judgements about **program** *effort, effectiveness, efficiency and adequacy*,
b. Based on *systematic data collection and analysis*,
c. Designed for **use** *in program management, external accountability, and future planning*,
d. Focuses especially on accessibility, acceptability, awareness, availability, comprehensiveness, continuity, integration, and cost of services.
(Attkisson & Broskowski, 1978: p. 24; emphases added)

1.3.3 Evaluation criteria in program evaluation

In assessing the value of a program, it is important to be clear about the criteria for judging value:

> "Social program evaluation is the process of thoroughly and critically reviewing the *efficiency, effectiveness* and *appropriateness* of any program or group of programs." (Senate Standing Committee on Social Welfare, 1979a, p. 5).

This definition has been widely influential in Australia (see Australian Council of Social Services, 1978; Department of Finance, 1994; Sharp, 1994) and is an important factor in the way many government and community agencies see and apply accountability and evaluation.

Each of these definitions will be elaborated below as they provide keys to introduce various important aspects of evaluation which are relevant to this review.

1.3.4 Disciplines of Inquiry: Research vs Evaluation

It seems that injury prevention (e.g., Gielen, 1992; Harrison, & Cripps, 1994) has to deal with at least four levels of interpretation:

- patterns of risk identified by the researchers and key opinion leaders in the community,
- the consequences of failure to guard against risk,
- the need for changing attitudes and behaviour of decision-makers and policy-makers, the community users of equipment, drugs, conveyances tools, etc., and manufactures of these products,

- the effectiveness, efficacy, efficiency and appropriateness of attempts to guard against risk.

All of these issues could benefit from increasingly systematic research and evaluation. The first two issues seem to be primarily the domain of researchers, the second two issues are more in the domain of evaluation.

Evaluation has been known under various names and forms in social sciences, education and social welfare fields. One important variant of evaluation which has been developed as its own field in education and health is *action research*. (e.g., Kemmis, 1985; Hugentobler, Israel & Schuman, 1992). For example, this comment in Scriven's (1991, p. 48) interesting *Evaluation Thesaurus:*

"ACTION RESEARCH 1. A little-known subfield in the social sciences that can be seen as a precursor of evaluation".

In health promotion evaluation is commonly found in "effectiveness studies" or "evidence-based intervention" (Macdonald, 1996), or "evidence-based medicine" (Sackett, Rosenberg, Gray, Haynes & Richardson, 1996).

Like the social sciences, the terms used in the field of evaluation are deceptively familiar. There is a need to continually recalibrate our usage of key terms in evaluation, and to remind decision-makers, as well as practitioners and consumers of evaluation, of the essential reason for clarity of meaning of often debased words. In this regard it is important for any review of the literature and annotated bibliography to clarify its usage of key terms (see also useful glossaries in Hawe, Degeling & Hall, 1990; Scriven, 1991).

Broadly speaking, although evaluations often use research methods as well as research related terms and frameworks, there is sufficient difference between the two approaches to warrant the specialisation in either field *per se*. Table 1 is an attempt to compare and contrast the common intended aspects and strengths and weaknesses of these two. Such gross comparisons often raise more conjecture than they solve. However, it is intended to provide a starting point for interpretation of the importance of evaluation as complementary to research approaches in injury surveillance and prevention.

TABLE 1

Comparison of Stereotypical Approaches (adapted from Isaac & Michael, 1995 & Patton, 1990)

	RESEARCH	EVALUATION
1. Purpose	New Knowledge, or Understanding; search for the "truth"	Achievement of "Mission" decide what is "useful"
2. Type of Questions Addressed	What is the cause? What is related to this? How widespread?	What is the result (intended or unintended)? How effective; efficient & appropriate? What should be done? How useful?
3. Assumptions	Causality is the key to understanding & valid theory The world is shaped by causal (cause and effect) relationships; Knowledge is important in its own right	Needs and stakeholder's interests or views are as important in decision making as knowledge or the "truth" depending on the assumptions and models of evaluation used (see Tables 2 & 3)
4. Dominant Types of Methods	Randomised Control Trials (RCTs) Experimentation Quantitative methods preferred Independent researchers Unobtrusive observational methods	Ranges from RCTs - through- Quasi-experimental designs - to - Participative inquiry Quantitative and qualitative methods Measurement is an intervention Case study
5. Intended Outcomes	Generalisable conclusions about process; Theory building; & understanding of relationships and predictability	Estimate of worth of process, output and outcomes Recommendations for change
6. Strengths	Greater control of plausible rival hypotheses; extraneous environmental factors and better interpretation of validity: internal (model/theoretical coherence) and external validity (generalisability)	Greater appeal to decision-makers Empowerment & ownership of stakeholders in developing decision-making process Accountability and Organisational learning
7. Weaknesses	Ethical concerns about control groups being withheld interventions External validity and timeliness out of kilter with needs of decision-making process	Threat to credibility of perceived political or stakeholder bias Internal or external validity may be weak
8. Preferred Meta-Analysis	Systematic reviews of comparable studies research designs (Randomised Control Trials) Peer review; Statistical synthesis of data across comparable studies, by comparisons of effect sizes	Comparison with accepted standards of evaluation ethics and practices (see Joint Committee on Standards 1994) How useful to decision-makers?

Much of the literature in the injury prevention and control field is concerned with research rather than evaluation. For example, consider the

"integrative review" paper which examined 15 experimental and quasi-experimental in the effectiveness of back injury prevention programs with the purpose of describing the state of knowledge in the field (Karas & Conrad, 1996). With the focus research, or on knowledge for its own sake, that paper was excluded from the present review and annotated bibliography, because the focus here is on literature concerned with evaluation (especially of injury prevention programs), ie., judgements about value of worth, not knowledge *per se*. However, the paper by Gielen (1992) is included despite its theoretical orientation, because it identifies a model for needs analysis and program planning. These are closely related complementary components to program evaluation.

1.3.5 Components of Program Evaluation

These terms will be explained in relation to a simple framework illustrating the relationship between them as illustrated in Figure 1, which uses the following concepts:

Needs:
predisposing factors (see Gielen, 1992; Green, Kreuter, Deeds, & Partridge, 1980; Green, & Anderson, 1986) or the identification of risk of injury, groups or individuals exposed to some risk, or the deficiency of information and/or lack of intervention to address an injury risk; in the community/policy arena (see also Baum, 1992; Hawe, Degeling & Hall, 1990; McKillip, 1987; Percy-Smith, 1996; Scriven, 1974, 1991; Siegel, Attkisson, & Carson, 1978;); this is sometimes seen as absolute need but more often relative need (`excess risk' etc.)

Goals:
a key link with strategic planning as a statement of intended outcomes (e.g., Attkisson, et al, 1978; Kiresuk et al, 1994), such as reduction of incidence of injury by 10% over the next triennium of program funding; this is the basis for identification and assessment of effectiveness (see Glossary in Hawe, Degeling & Hall, 1990);

Inputs:
the financial, physical and human resources which are allocated and consumed to *enable* a program or project to operate, e.g. the funds allocated to purchase safety equipment; the staff allocated to educate users, or maintain equipment (see also Department of Finance, 1994; Donovan, & Jackson, 1991);

Processes:
the operations of the intervention, or program being evaluated, which include the participation of the stakeholders in training, or operation of safety equipment, etc.; in the health promotion field these may also be *enabling* factors (see Auslander, 1996; Donovan, & Jackson, 1991; Gielen, 1992; Green et al. 1980; Green, & Anderson, 1986);

Outputs:
the products or immediate *results* created by the intervention or program (e.g. the number of trainees completing or graduating from safety training; the report of the number of cycles of safe use of the equipment, etc); most often used in productivity analyses (see Brinkerhoff & Dressler, 1990)

Outcomes:
the (more or less tangible) *consequences* for the stakeholders of the process and/or its outputs (e.g. skills gained in a safety training program; reduced incidence of accidents, safer work place, higher productivity of staff); in the health promotion field these may also be *reinforcing* factors which follow from the consequences of a particular intervention (see Auslander, 1996; Gielen, 1992; Green et al. 1980; Green, & Anderson, 1986; Hargreaves & Attkisson, 1978; Hudson, 1987; van Beurden, 1994); but they may also include unintended consequences.

A common fallacy is that outcomes are necessarily the impact of outputs (see Schick, 1996) when in practice they may be achieved through the selection or input stage or the process stages of the program (see Scriven, 1991; Schick, 1996; Sharp, 1994b; Sharp, Roffey & Lewis, 1993).

Any one evaluation may address more than one of these components, depending on the question at hand and the viewpoint. For example, "skills gained in a safety training program" may be an "outcome" for student evaluation of that program, but from the view of the manager, it is an output in terms of a wider safety program whose outcomes are to reduce injury occurrence/severity etc.

1.3.6 Criteria for Evaluation

The criteria for assessing programs in an evaluation usually consist of permutations and combinations of a number of client expectations or needs (Scriven, 1991), or program performance requirements or decision criteria (see Caulley, 1995; Owen, 1993). Here the following will suffice as they have become established as the main program evaluation criteria in several countries, especially the USA (Bruell, 1993; Scriven, 1991) and Australia (see Department of Finance, 1994; Hawe, Degeling & Hall, 1990; Sharp, 1994; Sharp, Roffey & Lewis, 1993):

Efficiency
is indicated by the number of outputs produced per input, or the amount of output (e.g. the specific tasks or products completed safely; or the number of graduates from a safety training program) for the given inputs (the resource consumed in $ and people costs). This is an important type of indicator in terms of accountability for the resources used, and productivity (see Brinkerhoff & Dressler, 1990) often used in occupational health and safety intervention evaluations;

Effectiveness
is an indicator of the extent that outcomes achieve objectives (i.e. did the project enable the participants to reduce their incidence of accidents; how many staff achieved the intended outcomes of the safety program). It shows the relationship between the outcomes for the intended recipients and the objectives of the project (see Brinkerhoff & Dressler, 1990; Hargreaves & Attkisson, 1978; Hawe, Degeling & Hall, 1990; Love, 1991). It helps demonstrate accountability for the performance of the project.

Efficacy
is the capacity for effectiveness under "ideal conditions" (or the implicit or explicit theoretical condition for effectiveness) such as in a laboratory experiment, or in a highly funded demonstration project vs effectiveness in real world applications (see Davis, 1992; Hawe, Degeling & Hall, 1990; Selby, 1994). Components of a program may be efficacious yet the program may not be effective

because of implementation faults (Hawe, Degeling & Hall, 1990). For example, it is reasonable to expect that helmets would reduce injuries for cyclists or seatbelts would reduce injury in road accidents, but the particular road accident prevention program may be ineffective because there was no compulsion to wear helmets or seatbelts, and seatbelts are effective only if they are worn properly. Indeed, in Australia, the Cyclists Rights Action Group (seehttp://www.pcug.org.au/~psvansch/crag/index.html) has opposed legislation requiring the wearing of helmets for various reasons concerning their efficacy (Curnow, 1995).

Appropriateness

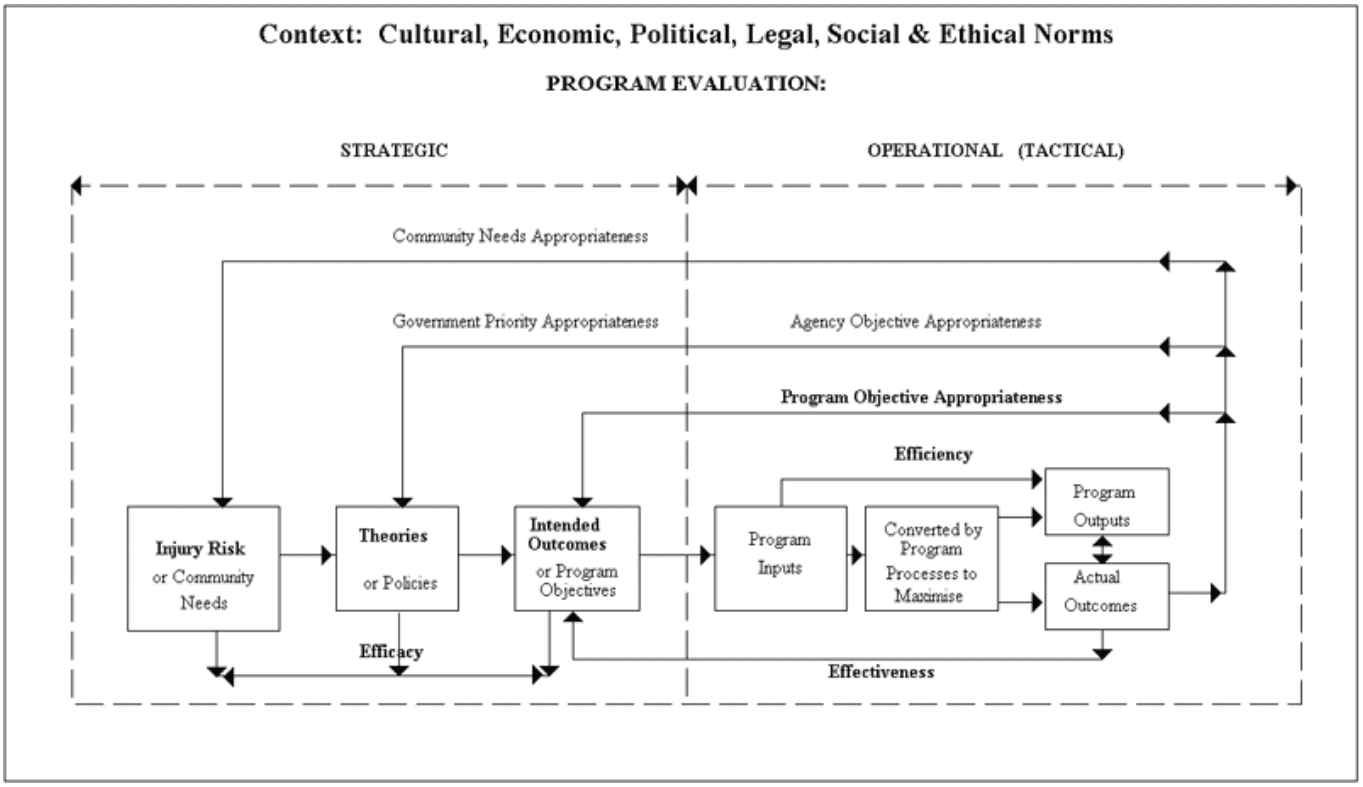
identifies the relevance of program objectives to community policies, or program participant's needs. This criteria addresses the broader social concerns and keeps a focus on the contribution of the program in terms of its context in society. It raises the question of unintended consequences. This can also be seen as an issue for effectiveness in the sense that "net effectiveness" may be less than "gross effectiveness", if unintended consequences or side effects are taken into account. Ethics and values issues are the keys to appropriateness evaluation (see Kimmel, 1988; Sharp, 1994). For example, an occupational health and safety program that requires the wearing of socially unacceptable or heavy and cumbersome protective gear may be efficient and effective, but is inappropriate to the social circumstances (e.g. hair nets for young men) or the physical environment (e.g., hot climate), and so causes unintended negative consequences like dermatoses, or increased risk of accidents due to discomfort and lapses of concentration.

Figure 1 gives an overview of the common relationships between these factors and the strategic versus operational processes of the program logic linking the program elements from needs or policy to outputs and outcomes.

There have been various attempts to identify or clarify evaluation criteria (see Caulley, 1995; Gross, 1979; Sharp, 1994a). Most commonly evaluation criteria include effectiveness and efficiency as the basic "values" for evaluation. However, another important evaluation criterion has often been ill-defined, viz: the term *appropriateness* in the context of evaluation (Sharp, 1994a; Sharp, Roffey and Lewis, 1993). In an attempt to better define appropriateness, a survey of Australian Commonwealth and State government documents and found at least five connotations of the term appropriateness can be distinguished (Sharp, Roffey and Lewis, 1993), namely:

- **strategic** aspects (in planning and management processes, as distinct from the "efficiency and effectiveness" operational focus of traditional evaluative criteria); strategic aspects usually refer to the degree to which the program as a whole relates to its political, economic, legal and socio-cultural environment;
- **coherence**, concerning the logical consistency of the program elements as evaluated;
- **usefulness**, sometimes referred to as relevance, or feasibility, and often a critical aspect of the evaluation of the quality of a program's design, structure, and implementation;
- **probity**, goodness and integrity with reference to cultural, political, social justice, social and ethical norms.
- **expediency**, when the documents did not define or clarify the meaning of appropriateness, or used in a manner which gave a conveniently vague or political scope, this was called "expedient".

FIGURE 1: Relationship between Evaluative Criteria



This diagram (based on Sharp, Roffey & Lewis 1993; and Sharp, 1994a) gives a perspective on the relationship between the four main evaluation criteria used here in evaluation: appropriateness, efficacy, efficiency & effectiveness

The *Human Services Thesaurus* (Department of Human Services and Health, 1995) does not deal adequately with the three criteria of program evaluation used in the Commonwealth (cf. Sharp, 1994). The *Human Services Thesaurus* does list effectiveness and efficiency, but not appropriateness.

Other criteria often used in evaluation include: economy and cost-effectiveness (see Caulley, 1995; Gross, 1979; Owen, 1993).

According to the Cochrane Collaboration effectiveness and efficiency are key evaluation criteria, but in the spirit of Cochrane's (1972) work appropriateness may be as important in medicine as in other human services. Indeed, ethical considerations are very much about appropriateness (Kimmel, 1988; Sharp, 1994a).

1.3.7 Qualitative vs Quantitative data collection and analysis

There has been a long running debate (see Lincoln & Guba, 1985) about the relative strengths and weaknesses of quantitative approaches (see Brinkerhoff & Dressler, 1990; Gruenewald, Treno, Taff & Klitzner, 1997) and qualitative approaches (see Patton, 1987; 1990) to evaluation[8]. It is now generally accepted that both approaches ought to be included in evaluation designs in order to triangulate data (Dennis, Fetterman & Sechrest, 1994; Patton, 1990). Indeed it is appropriate to cross-validate (triangulate) data with alternate methods and increase the degree of inference about relationships (de Vries, et al., 1992; Love, 1991; Patton, 1990; 1997; Steckler, et al, 1992).

Dennis, Fetterman and Sechrest, (1994) pointed out that despite the rhetoric about integrating the two approaches there are seldom systematic and coherent examples of this in practice. They identify useful strategies of how to integrate qualitative and quantitative approaches in the design, implementation and reporting phases of evaluations of substance abuse prevention or treatment. There are many lessons which they refer to that have relevance to injury prevention evaluation.

---

[8] See for example the excellent Sage Publications series on *Quantitative Applications in the Social Sciences* and the Sage series on *Qualitative Research Methods*

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

1.4 EVALUATION FORMS

One of the more informative reviews of the evaluation literature is by Owen (1993). He pointed out that in addition to models (see below), evaluations may take on various forms and approaches (e.g., Owen, 1993) depending on the emphasis on different aspects of the program or the decision-making involved. Table 2 identifies some of these considerations.

TABLE 2: Forms of Evaluation
(based on Owen, 1993, p.22)

Form	Key Terms, Issues and Questions
Evaluation for **Development**	• needs analysis and needs assessment • ex ante evaluation or planning before the program is designed What are the appropriate objectives to meet the needs? What costs/benefits are able to be achieved?
Design Evaluation	Clarification of "program logic" through: • Underlying structure of inference of causal relationships between interventions and performance • Program rationale; • Evaluability assessment (see Davis & Salasin, 1987; Wholey, 1994). How does the program and/or evaluation design provide strength of inference of linkages and causal relationships between the intervention and the results? What are the intended outcomes and how can they be measured? What performance indicators are going to be useful?
Evaluation in Program **Management**	Can be "formative" or "summative" evaluation (See Scriven, 1991) Formative is an ongoing program improvement role while summative is a cumulative examination endpoint in time or key milestone. Accountability issues are often addressed. How much can management do to improve the program?
Process Evaluation	Often "formative evaluation", related to quality issues through: • Monitoring or collecting information about program interventions, event, and activities • Making decisions during program implementation • Advising program deliverers or managers about improvements in their practice; How can stakeholders better understand and improve the processes or interventions and the relationship to performance the program?
Impact Evaluation	"Summative evaluation" What has been the impact of the intervention process? Were the intended outcomes achieved? Were there unintended outcomes? What was the overall worth of the program?

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

1.5 MODELS OF EVALUATION

In order to classify the evaluation literature it is also useful to refer to the types of models (Caulley, 1989; Herman, Morris & Fitz-Gibbon, 1987) and the forms of evaluation (Owen, 1993). Others have suggested that it is possible to identify "paradigms" (e.g., Smith & Glass, 1987; Lincoln, 1992), but the field is still evolving and too diverse for such interpretation.

Table 3 outlines some of the critical features of various evaluation models.

TABLE 3: Some Models of Evaluation (based on Caulley, 1989; & Herman, Morris & Fitz-Gibbon, 1987; Isaac & Michael, 1995)

TYPE OF MODEL	KEY FEATURES & ISSUES	KEY AUTHORS
A. Measurement-Oriented	Monitoring through measurement of outputs and/or outcomes or attributes of the target population Sometimes referred to as the earliest form of evaluation - but long since differentiated from program evaluation per se. Using consistent performance indicators in monitoring is often a necessary but not sufficient basis for evaluation. (e.g. there is much more to benchmarking than establishing quantitative performance indicators).	Most recently has re-emerged in the charting of Government service delivery indicators (1997) and the performance indicators debate (cf. Bartos, 1995; Winston, 1995) **Benchmarking** has emerged as a management oriented external comparative evaluation using industry-based performance information (cf. Camp, 1989; Sharp, 1994)
B. Objectives, Goals - Oriented	To clarify purpose, and achievement, goals (and the monitoring of attainment of goals) are seen as essential bases for evaluation, especially of effectiveness	Many evaluation authors & management paradigms follow this model, e.g. Attkisson, et al, (1978) Kiresuk et al (1994); Drucker (1973)
C. Needs-oriented or Goal-free	Needs analysis and assessment has always been an important part of ex ante evaluation, (ie., prior to program beginning) usually in clarifying goals or in cost-benefit analysis Independently of managers' or program	e.g. Siegel et al (1978) Percy-Smith (1996) Scriven (1974; 1991) has adopted a hard line on the importance of independence from pre-conceived goals

	goals, evaluation should focus on the clients' needs and whether they are met.	
D. Decision-making or Management-oriented	The critical role of evaluation is to provide a basis for better management through providing information to assist in decision-making	e.g Stufflebeam's (1971) CIPP model (see below)
E. Experimental or evaluation research	Focusing on experimental and quasi-experimental designs so that evaluation can explain and generalise the results of programs through assessing for cause-effect relationships.	Campbell (1969) saw social reform programs as experiments in policy development (see also Cook & Campbell, 1979; Weiss 1972) "What works?" questions in public health tend to focus on this model.
F. Expert Opinion, Panel or Commission	Often in highly charged circumstances where a parliamentary review or evaluation is required - an expert Commissioner or panel of experts are convened to conduct the evaluation	e.g. the Baume inquiry (Baume, 1991; Senate Standing committee on Social Welfare, 1979) in Australia; or the Gore (see Bruell, 1994) review of administration in the USA. See also Meister (1989) on ergonomics.
G. Self-Evaluation, or Empowerment evaluation	Usually Formative evaluation conducted by the manager or staff of a program for their own purposes Contrary to the conventional scientific or "positivist" paradigm Empowerment of the stakeholders using evaluation as part of the advocacy or self-determination process by groups of stakeholders	e.g. see Kemmis' (1985) approach based on action research and action learning principles; (see also Wadsworth, 1984; 1991) Guba & Lincoln (1989) have coined the phrase Fourth Generation Evaluation for their "constructivist" approach see Fetterman (1995)
H. Responsive or Democratic	Evaluation is to assist to clarify the program's values, processes and stakeholders' perspectives	Stake's (e.g. 1980) "responsive evaluation"
I. Utilization-focused	A situational approach to focus on the intended use of the evaluation by the stakeholders, to produce an evaluation which is used by them	Patton (1997)

Broadly speaking one could identify models A to E as following the more conventional, "scientific" or "positivist" approaches to inquiry, where independence and experimental design are the keys to validity and interpretability of evaluation. Models G, H and I are often at odds with the pretensions of "objectivity" of the models A to E. Models G, H and I are emerging as gaining stronger interest in evaluation (Lincoln, 1992; Patton, 1997). The latter models expect the evaluator(s) to be deliberately participative in engaging with the program managers and other stakeholders. Model F could adopt either the "positivist" or "constructivist" approach, depending on the appointed expert(s).

Below we will briefly introduce models A, B, C, E, F, G and I to give a basis for analysis of the injury prevention evaluations reviewed. These models were chosen because of their common currency in the related fields of human services and health.

1.5.1 Measurement Model or Benchmarking (Model A)

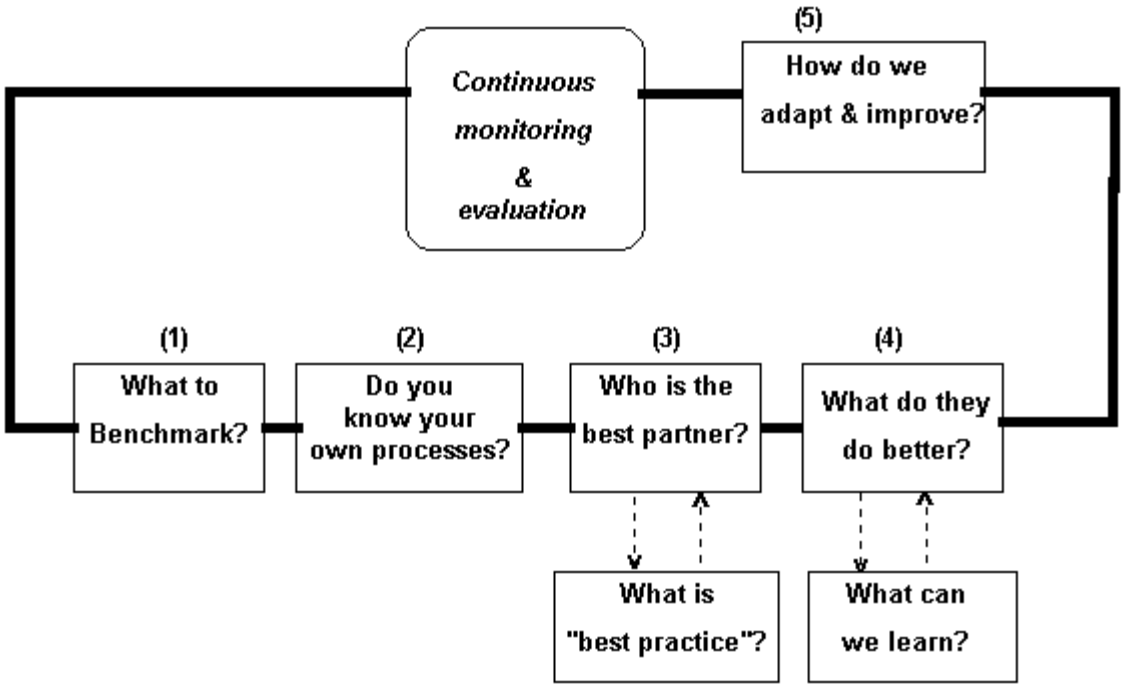
Social indicators of need and community indicators of the impact of social program interventions have many applications (e.g., Gruenewald, Treno, Taff & Klitzner, 1997). Similarly, there is a long history in occupational health and safety research and evaluation of focusing on such measures or indicators as *productivity* (Brinkerhoff, & Dressler, 1990).

As a form of measurement-oriented model benchmarking is an extremely important strategic approach to evaluation which has received a lot of interest (see Bartos, 1995; Sedgewick, 1995; Sharp, 1994b).

The term "benchmark" generally refers to the collection of external comparative quantitative data which can be used in organisational diagnosis to inform management (Harrison, 1994; Schubert, 1997). The focus is supposed

to be on "best practice" as the benchmark to aim for in assessing organisation effectiveness against the external environment. "Benchmarking", however, is more problematic as there are various approaches advocated from a variety of "experts". Generally, *benchmarking* refers to the management of the change process in organisations which uses benchmarks, and comparisons between the organisation in question and exemplars of "best practice" (see Sharp, 1994b). Most introductions to benchmarking identify five or seven steps. However, Sharp (1994b) pointed out that there are at least a few more unstated processes, hence "the magical 7 steps (plus or minus 2)" provided by figure 2.

Figure 2
7 STEPS TO BENCHMARKING
(from Talbot & Sharp, 1994 adapted from Sharp, 1994a)



An example of a benchmarking strategy in Government is the Council of Australian Governments Steering Committee for the Review of Commonwealth/State Service Provision (1997) which has developed a framework and gathered data on the performance indicators of the effectiveness efficiency and appropriateness of the provision of community services by Governments. Such databases and performance monitoring are intended to improve the evaluation and benchmarking of Government service provision.

The key informants identified the validity and reliability of data sources and the comparability of indicators of injury as common problems in injury prevention research and evaluation. One area which most agreed could be a useful investment in the field is the development of standardised approaches to data collection and reporting of injury data. Such an approach is already being applied in other areas of human services (see Industry Commission, 1997)

In Australia in the health promotion field, a similar approach is being developed by van Beurden (1994), with HOOPS (Health Outcome Oriented Problem Segmentation). This is a pro forma approach to program logic analysis and data collection, with a graphical display framework for health promotion planning, program development and evaluation.

Some health outcome indicators have long been used more-or-less in this way. Infant mortality rates are perhaps the best example.

1.5.2 Objectives, Goals-Oriented Model (B)

The usefulness and credibility of program budgeting (or program planning and budgeting system), management by objectives (MBO, Drucker, 1973) and other methods of accountability in management depend on these approaches being linked to the planning and goal formation within the organisation or project (Hatry, 1990; Hudson, 1987; Kiresuk, Smith & Cardillo, 1994; Love, 1991; Wholey et al., 1994). In order to check on how well the

program or project is achieving its goals (i.e., how effective it is) requires some measures or indicators of the performance of the program or project towards its goals. Some methods for monitoring performance towards goals which have been useful in the human services are: *Global Assessment Scaling* and *Goal Attainment Scaling* (see Hargreaves & Attkisson, 1978; Kiresuk & Lund, 1978; 1994; Kiresuk, Smith & Cardillo, 1994; Sharp, 1997b).

However, it is wise to incorporate checks and comparisons which build validity in the plan of an evaluation. This can be done by using alternative sources of information, through development of various measurement techniques. In other words it would be advisable to also apply more conventional techniques for performance planning and measurement, as well as goal attainment scaling.

Goal Attainment Scaling

In the complex field of injury prevention it is often necessary to justify programs and evaluate them based on specified legislative or policy goals. It is sometimes difficult to adequately describe the expected or desired outcomes in measurable terms. One technique which is useful to both clarify the objectives, the indicators and the evaluation of a new program is called "Goal Attainment Scaling" (GAS).

Basically the strength of the GAS technique is that it accommodates both quantitative and qualitative data about the performance of a program in terms of the goals of the participants. It engages all key stakeholders involved in an evaluation to form an agreement about the most important aspects of the expected, desirable and undesirable outcomes of the program upon which they are about to embark. By providing a common format for the expression of the goals and/or objectives, the technique forces participants to document in an unambiguous way the expected outcomes, as well as the less than expected and much less than expected results. GAS provides a common scoring framework which enables comparative data collection and statistical manipulation across individuals and groups of participants. During the program and when the program is completed the stakeholders are accountable for what they have done, through GAS they can at least have a common framework for estimating the benefits, and general (intended and unintended) effects of the program.

This technique is cursorily illustrated in Figure 6, which provides a fictitious example of Goal Attainment Scaling applied to the prevention of youth accidental injury.

1.5.3 Needs-Oriented or Goal Free Model (C)

This approach addresses the actual outcomes in terms of the expectations or needs of the end-recipients of the program. The model deliberately attempts to avoid the intended processes or outputs of the program as being biased by the managers' point of view. Going directly to the target clients or end recipients of the program or intervention, is a controversial approach (Scriven, 1972; 1974; 1991). The attempt to be independent of the managers' statements of policy and goals, is obviously difficult if they are the fundors of the evaluation.

1.5.4 Decision-Making / Management Oriented (D)

Decision research has emerged as a field of its own (see Carroll & Johnson, 1990). But evaluation has frequently been seen as a tool of decision making. According to the approach to decision-making advocated by Stufflebeam (Stufflebeam 1971; Stufflebeam, et al, 1971). Like Figure 1, this model emphasises that the evaluation should include analyses of the 4 CIPP factors, viz:

- context (e.g. political, economic and social factors, internal and external to the program)
- inputs (the resources such as staff, assets and cash)
- processes (the operational elements involving the delivery of the program)
- products (the results including output and outcomes)

This model is also one of the bases of the classification of forms of evaluation outlined in Table 3.

1.5.5 Experimental or evaluation research (Model E)

The closest to the positivist orientation focusing on deriving cause and effect inferences from the rigorous design of

the program and its evaluation (see e.g., Cook & Campbell, 1979; Scriven, 1991). Many texts on evaluation concentrate on this model (e.g. Rossi and Freeman, 1989) However, common practice in evaluation often fall short of the scientific aspirations of this model.

Nevertheless, this is obviously the strongest form of evaluation with relevance to the epidemiology (e.g., Sackett, et al., 1985), and "evidence-based medicine" (Sackett, et al, 1996) approaches to injury prevention.

1.5.6 Expert Opinion, Panel or Commission (Model F)

The Expert Opinion or "Connoisseurial Assessment" approach (see Rossi & Freeman, 1989, pp 364 -370) is seldom developed in the literature nor critically reviewed (see Scriven, 1991). However, the discipline of *Ergonomics* has made a profession of the expert evaluation of programs and injury prevention techniques in the occupational health field (see Granjean, 1980; Meister, 1986).

For example, a very inexpensive and important evaluation technique in occupational health would be the "walk-through" (see Meister, 1986, pp. 59- 61). This basically entails an initial "mock-up" of the occupational site, with the injury prevention program or intervention in place, to test its feasibility or efficacy under controlled (artificial) conditions. Then when the mock-up works satisfactorily, the injury prevention intervention can be installed in the real workplace and after an onsite inspection or walk-through, the program begun. A subsequent walkthrough to monitor or re-evaluate the site would be the basis for an expert opinion report on the evaluation of the injury prevention intervention's efficacy.

1.5.7 Self - Evaluation and Empowerment Evaluation (Model G)

One of the latest uses evaluation to gain recognition is its ability to empower the stakeholders in the decision or advocacy process (Everitt, 1996; King, 1996; Sharp, 1995; Wadsworth, 1996). According to Fetterman (1995) this is worthy of a designation of a new field of evaluation, which he calls: *Empowerment Evaluation*.

Here there is a clear commitment to using evaluation as a tool for stakeholder group self-determination and control or at least major participation in the evaluation and strategic decision making. It is basically self-evaluation practiced by semi-autonomous groups.

Fetterman's (1995) basic steps towards Empowerment Evaluation can be identified as:

a. stakeholders meet to examine the current and future scenarios - and the influence of internal and external forces on the program (including a SWOT analysis);

b. stakeholders agree on common goals for program improvement;

c. stakeholders meet to determine the relevant strategies to achieve the proposed program goals and intended outcomes;

d. evaluation advisers assist program stakeholders to determine the type of data collection methods relevant to gather and analyse the relevant data for their evaluation and decision making about the program.

Fetterman (1995) points out that an important ingredient for all this to happen is that the stakeholders should have a committed dynamic core of honest self-critical and supportive members who are willing to experiment and learn by trial and error.

In injury surveillance and prevention this approach is possibly relevant to the industrial context (cf. Burdorf, et al., 1997; Zwerling, et al. 1997). In that field there is a history of advocacy related research and evaluation of work related risk factors (see De Cock; 1986; Gardner & Palmer, 1992; Meister, 1986; Williams, 1986).

1.5.8 Responsive or Democratic (Model H)

There is insufficient difference between this model and models G and I, these days. So it is not elaborated here (see Scriven, 1991; Stake, 1980).

1.5.9 Utilization Focused Evaluation (Model I)

The focus of this popular approach is on the needs of the decision-makers and how to meet these needs in maximising the usefulness of the evaluation.

Indeed, usefulness and utilisation are common themes or credos of program evaluation. There are a number of models of evaluation which address these issues. But the approach most notably relevant is that identified with Michael Quinn Patton (1997) author of *Utilization Focused Evaluation.* As he points out:

> "How evaluations are used affects the spending of billions of dollars to fight problems of poverty, disease, ignorance, joblessness, mental anguish, crime, hunger, and inequality. How are programs that combat these societal ills to be judged! How does one distinguish effective from ineffective programs! And how can evaluations be conducted in ways that lead to use? How do we avoid producing reports that gather dust on bookshelves, unread and unused?"
> (Patton, 1997, p. 4)

Recent approaches to evaluation in the field of injury prevention (see e.g., Danneberg, 1996; Miller, 1996; Towner, 1996; Zwerling, Daltroy, Fine, Johnston, Melius,. &. Silverstein, 1997) seem to be predicated on such assumptions as the following:

- evaluations are concerned with providing useful information for decision-makers
- evaluations are to assist the researchers in clarifying the causality of models of risk factors and preventative factors
- quantitative data is more convincing than qualitative data
- controlled quasi-experimental designs are more powerful means of evaluation than other designs

But as Patton (1997) points out:

> "There is no one best way to conduct an evaluation. ...
> The design of a particular evaluation depends on the people involved and their situation. ... The standards and principles of evaluation ... provide overall direction, a foundation of ethical guidance, and a commitment to professional competence and integrity, but there are no absolute rules an evaluator can follow to know exactly what to do with specific users in a particular situation. ... This means *negotiating* the evaluation's intended and expected uses.
> Every evaluation situation is unique. A successful evaluation (one that is useful, practical, ethical and accurate) emerges from the special characteristics and conditions of a particular situation - a mixture of people, politics, history, context, resources, constraints, values, needs, interests, and chance. ...
> The right way, from a utilization-focused perspective, is the way that will be meaningful and useful to the specific evaluators and intended users involved, and finding that way requires interaction, negotiation, and situational analysis." (Patton, 19978, p. 126)

Essentially, the approach advocated in this literature review is pragmatic view, which follows the "situational evaluation" path.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

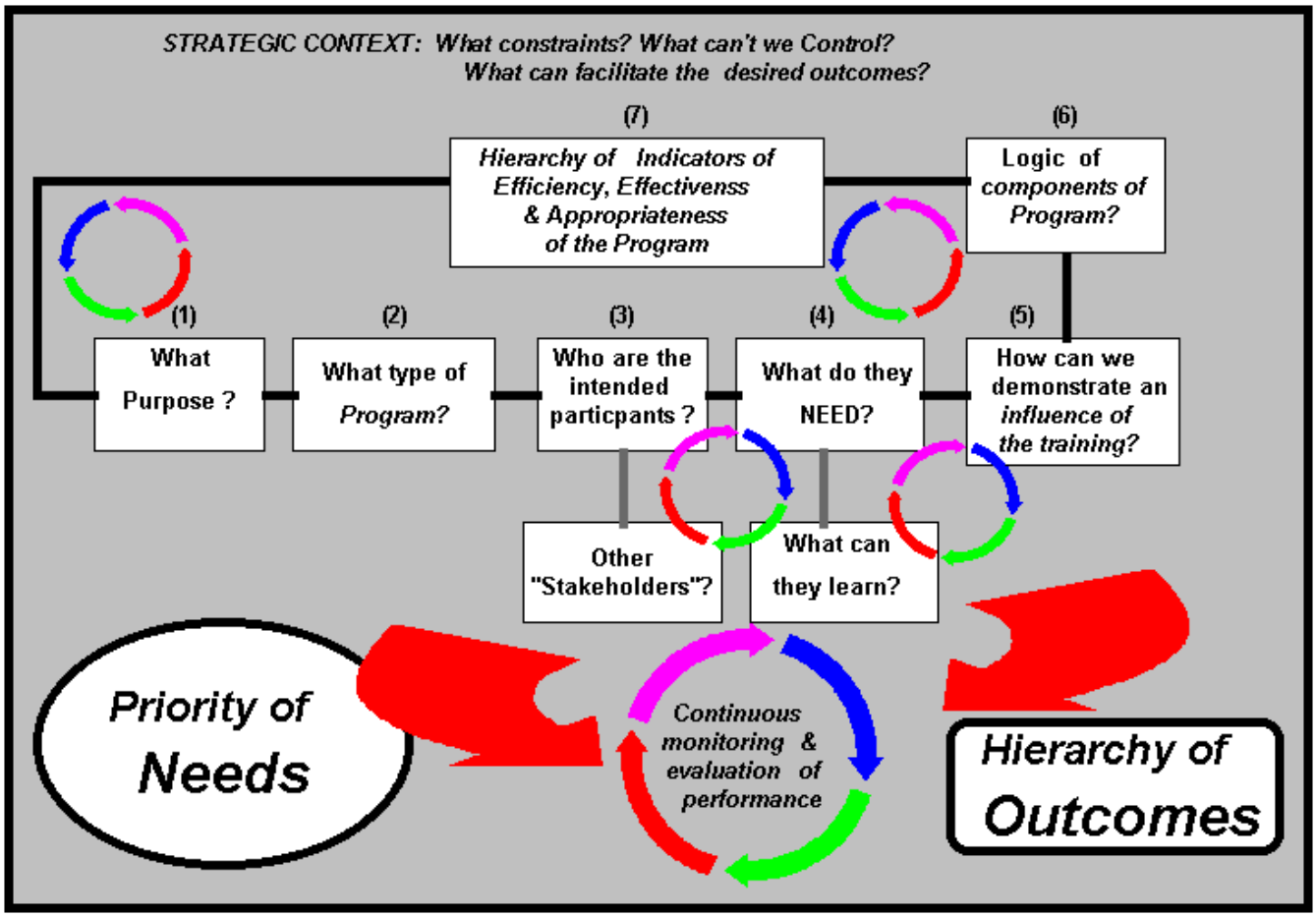
1.6 PROGRAM PLANNING & EVALUATION DESIGN LOGIC

Most writers on evaluation emphasise the importance of preparation and planning for evaluation as part of program planning and design (e.g. Davis & Salasin, 1978; Rossi & Freeman, 1989; Winston, 1993; and Wholey, 1994). For most approaches to program management it is important to establish goals and objectives in the context of the injury prevention strategy and management policies. A technique often used in program evaluation to link the evaluation of a program with the total cycle of the program's planning, initiation, operation, outcomes, impact and decision processes is called "*program logic*" or implementation analysis. Program logic is used to analyse the overall framework of objectives into a logical causal relationship (see Lenne & Cleland, 1987). Generic outcome hierarchies can be developed which can be used as templates or guides to review the training program's implementation and check the results obtained in terms of the intended outcomes.

From the analysis of program logic it is possible to identify the timing and linkages in the sequence of events and cause and effect relations (see Kelly & McGrath, 1988). Indeed, the basic approach is akin to the "evidence diagrams" (e.g. Rivara, et al. 1997) often provided in association with regression analysis research.

Here the concept of program logic is introduced through a series of steps in an overview scheme of evaluation as a management and accountability system (see Figure 3). Then Section 2 will elaborate on the types of tools (see Charts I to VI) for implementing this scheme in terms of the linkages between the design of the injury prevention program, or control intervention, and the evaluation of it.

FIGURE 3: The Basic Steps of PROGRAM LOGIC ANALYSIS & EVALUATION



1.6.1 Basic Steps in a System of Planning & Evaluation

In essence, planning and evaluation form the processes of feedforward and feedback which are fundamental engines in any management system (see Checkland, 1981; Checkland & Scholes, 1990; Leeuw, Rist, & Sonnichsen, 1994; Mason, & Mitroff, 1981; Senge, 1990; Sharp, 1997a; Stacey, 1993a 1993b; 1993c; Stata, 1989).

A useful link between planning and evaluation comes from the **program logic** method of evaluation design, which involves the following steps:

- clarify the **(1) purpose (including aims/goals)** and **(2) type** of the program being reviewed whether it is focused on intervention or prevention or training per se, or more broadly educational, or advisory; versus an organisational culture or attitudinal change management program or on environmental change or regulatory change etc. (see the process in Figure 3 and a fictitious example of the product of Figure 3 in Figure 4);
- identify (3) who are the **intended participants** and describe characteristics relevant to subject (e.g., level of risk, demography, numbers and types of injuries prevalent);
- what do they **(4) need**? (e.g., analysis of risk, reduce exposure, increase knowledge, reduce incidence of injury) - *needs analysis* is a significant field in its own right which cannot be elaborated upon here (see McKillip, 1987; Percy-Smith, 1996; Siegel, Attkisson & Carson, 1978), but which will form an important basis for the conceptualisation of the evaluation and clarification of the methodology (see Charts II and III below)
- how can we demonstrate (5) that the injury prevention program or the environmental characteristics etc. **influenced the participants** and contributed to the reduction in need, or satisfaction of desired outcome?
- how can we determine whether any differences obtained were intended or unintended? This is an efficacy issue, pertaining to the investigation of cause and effect relationships, which require sophistication of the logic of the design of the program and the evaluation (see Chart I below).
- determine the **(6) logic of the operational relationships** between the components of the policy, objectives and operations of the program. Can we show the chain of inference or deduction between stages or components of the program or system in its implementation (see the processes in Figure 3 and a fictitious example in Figure 4 and follow the steps in Chart IV).
- develop and use a **(7) relevant outcomes hierarchy** to generate specific indicators or statements of outcomes for the program (e.g., specify attributes of the stakeholders and what their expectations are for what type of results, etc.); specify *efficiency, effectiveness (efficacy) and appropriateness* indicators in terms of program inputs, process, outputs and expected outcomes (see a fictitious example in Figure 5).

Above all there must be continuing interactive processes of monitoring and evaluation of performance of the participants and the program, e.g., use *benchmarking* (see Sharp, 1994) to compare the outcomes generated by the program with those from an outcomes hierarchy used in comparable organisations, or programs, to see whether there are major differences and to interpret why such differences might occur;

a. decide whether any of these differences are unsatisfactory and take appropriate steps to rectify;
b. identify what factors would have contributed to the successful achievement of these outcomes (by consultation with key informants and other stakeholders, and by reviewing the implementation processes conducted);
c. which of these factors were due to the operation of your program and can be demonstrated to be controllable?
d. which of these factors were not due to the operation of your program or did not appear to be controllable?
e. determine whether these factors can be brought under managerial control in the program;
f. what are the activities which can operate to produce success factors in future?
g. examine the program and outcomes again in terms of the generic hierarchy and determine the level of information required to monitor and manage the process in future.

Fictitious Example

The techniques of program logic and Goal Attainment Scaling are illustrated in Figures 4, 5 and 6, which deal with a fictitious example of how to clarify the objectives intended outcomes and performance indicators of a supposed program seeking to provide information advocating for the prevention of accidental injury in young people (especially males aged 15 to 20). For purely a practical demonstration of the kinds of material which could be developed for injury prevention programs in the process of preparation for evaluation, the following goals are considered for an imaginary injury prevention program, say being considered for funding or evaluation by a state health department, targeted at the 15 to 20 year old males:

Goal 1: The injury prevention program will increase awareness of the extent and dimensions of youth injury (especially in males aged 15 to 20 yrs) over school holidays
Goal 2: The injury prevention program will be associated with a decreased incidence of severe injury in males aged 15 to 20 yrs by next budget

These goal statements could be re-stated in more precise terms, but they will do for the present purposes.

FIGURE 4
Analysis of the Suggested Needs & possible Risks underlying an Injury Prevention Program

Let us imagine the following Goals for a State Government Health Department:
Goal 1: Increased awareness of the extent and dimensions of youth injury (especially in males aged 15 to 20)
Goal 2: Reduced incidence of severe injury in males aged 15 to 20

The Implications of the terms and intent of Goal 1

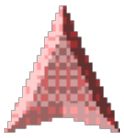
CONCEPT	MEANING	INDICATOR(S)
increased	a demonstrated *change* of some indicator(s) of awareness	• upwards trend from baseline data (e.g. demand for and distribution of safety literature); • improved quality (e.g. more targeted school and sport safety literature)
awareness	knowledge (i.e., data access and meaning) targeted safety literature *by Whom?:* Politicians (e.g., State; Commonwealth; Local Councillors) collectives and advocates (e.g., Youth Councils; Sporting Clubs; Schools & affiliated programs); individuals (e.g., unemployed young people; participants in sporting programs)	• content and quantity of media reports; • letters to press, politicians • content and quantity of advocacy; • content and quantity of conference papers; • demand for literature through mailing lists
extent of youth injury	*felt need* => models of effects of sports injury; risk/incidence of injury in young males aged 15 - 20	• consensus of young people in focus groups; public forums (e.g., Youth Festivals) • hospital admissions of young males 15 - 20
dimensions of youth injury	qualitative (e.g., social/health consequences of injury); quantitative (e.g., demographic, economic)	• individual case studies of social/health consequences of injury); • surveys of hospitals and GPs, Sporting clubs & school Physical education teachers

---

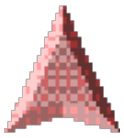
FIGURE 5

Outcomes Hierarchy for Goal 1

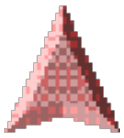
7. Increased awareness of need for and Opportunity for targeted youth injury prevention programs



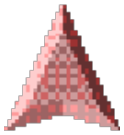
6. *Young people* have better **access to and make use of higher quality targeted safety/injury prevention programs**



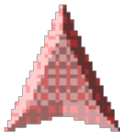
5. *Researchers and youth workers* have better **access to and make use of higher quality targeted safety/injury prevention programs**



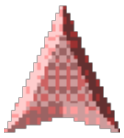
4. *Injury Prevention projects & safety advocates* improve the **funding strategies, targeting and delivery of their advice, research and services.**



3. *State Health Agencies* are better able to service and equip *Policy Makers and Safety Advocates* with necessary **knowledge on how to target safety/injury prevention programs.**



2. New models of injury risk and injury prevention and control **are communicated** by *State Health Agency* to *Government, news & professional media*



1. Advances are made in **models (extent & dimensions) of youth injury** by *State Health Agency.*

FIGURE 6: GOAL ATTAINMENT SCALE GOALS 1 & 2

Level of Expected OUTCOME		**Behavioural Statements of EXPECTED OUTCOMES**	
Rating		Goal 1	Goal 2
MUCH *MORE* than EXPECTED	**+2**	Over 80% of Schools and >70% of the Sporting Clubs request information on Youth injury risk AND Politicians propose to introduce incentives for preventative programs and/or reduce sales tax on protective equipment	About 90% of surveyed GPs and Hospitals report reduced Youth accident related injury AND/OR *Large* (20+%) & Statistically (p< 0.01) significant **reduction** in next year's incidence of accidental death of males 15 - 20 yrs
MORE than Expected	**+1**	About 75% of Schools and/or about 50% of the Sporting Clubs request information on Youth injury risk AND/OR Politicians raise the need for preventative programs in Parliament	About 60% of surveyed GPs and Hospitals report reduced Youth accident related injury (others report static rates) AND/OR Statistically (p< 0.05) significant reduction in next year's incidence of accidental death of males 15 - 20 yrs
EXPECTED	**0**	About 50% of Schools or about 30% of the Sporting	About 40% of surveyed GPs and Hospitals report

Outcome		Clubs request information on Youth injury risk OR Politicians request information on preventative programs and/or incidence of youth injury	reduced Youth accident related injury (others report static rates) AND/OR Some (non-significant) reduction in next year's incidence of accidental death of males 15 - 20 yrs
LESS than Expected	**-1**	About 25% of Schools or about 10% of the Sporting Clubs request information on Youth injury risk OR Politicians **do not** raise the need for preventative programs in Parliament	About 30% of surveyed GPs and Hospitals report reduced Youth accident related injury (others report increased rates) AND/OR No reduction in next year's incidence of accidental death of males 15 - 20 yrs
MUCH *LESS* than EXPECTED	**-2**	Some Schools and Sporting Clubs complain that information on Youth injury risk is unnecessary or alarmist AND/OR Politicians oppose the need for preventative programs in Parliament	Less than 25% of surveyed GPs and Hospitals report reduced Youth accident related injury (others report increased rates) AND/OR *Large* &/or Statistically ($p < 0.05$) significant **increase** in next year's incidence of accidental death of males 15 - 20 yrs

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

2.1 DECIDING HOW TO DESIGN AN EVALUATION

A numbers of leaders in the field of program evaluation have pointed out that there is often more art and individualism to the planning and design of program evaluation studies than science and standards (e.g., Cronbach, 1982; Patton, 1990; 1997; Rossi & Freeman, 1987).

Depending on the matter to be evaluated, and the aims of a particular evaluation of it, any of a wide range of evaluation methods and study designs may be appropriate. To assist in the decisions about designing an evaluation, a series of flow charts are provided to address various questions and decision points (see Figure 7, and Charts I to VI).

These charts and the associated annotated bibliography are not intended to 'automate' the process of deciding which approach to evaluation design should be adopted in any particular instance. This remains a matter for professional judgement, negotiation with stakeholders, and so on. The charts are intended to assist this decision process by providing a conceptual framework for evaluation, a lead into useful texts on evaluation practice, and a guide to examples from the literature, especially the injury prevention literature. The conceptual framework is intended to show the breadth of the range of approaches that are taken to evaluation, and to make it easier for practitioners to think about the range of approaches. The examples from the literature (especially the injury prevention literature) are intended to give a practical flavour to what might otherwise be an abstract subject. They are a selection from the literature available from our search, but are not intended to be comprehensive.

This section introduces the series of charts, shows how they were developed from the literature and provides brief comments on some considerations that may be pertinent when using the framework with examples to illustrate how to plan a particular evaluation.

These comments do not amount to a systematic treatment of the process of selecting an evaluation method, and interested readers should refer to the various handbooks which make a feature of this decision process (e.g., Herman, et al., 1988; Owen, 1994; Rossi & Freeman, 1989; Wholey et al., 1994).

On the basis that a proposal for an intervention has been decided Figure 7 asks the key question: "Will X be evaluated?" This is an important issue often glossed over in the early stages of planning an intervention program. Of course Planners and Managers will accept that evaluation is relevant, at *some stage.* But we insist, along with many others (see e.g., Herman, et al., 1988; Isaac & Michael, 1995; Owen, 1994; Rossi & Freeman, 1989; Wholey et al., 1994), that evaluation must be taken seriously at this point. If the answer to the first question in Figure 7 is a hesitation or resistance, then the accountability mechanisms, funding agreements, appointments of staff, all should be reviewed in the context of the uncertainty that lack of evaluation planning will bring.

Once it is agreed to plan the evaluation, program logic (see Figures 3, 4 and 5) and Goal Attainment Scaling (see Figure 6) and other tools could be used to clarify the aims of the program, and of the stakeholders which will provide a foundation for the evaluation. The various problems which could impact of the clarity of the plans for the program and the evaluation are detailed in Charts II [PDF, approx 20 Kb] and III [PDF, approx 31 Kb], along with the background details on the models of the injury intervention (Chart V [PDF, approx 18 Kb] and VI [PDF, approx 32 Kb]).

But first and foremost it is appropriate to followup Figure 7 with three key questions (which are the bases of Chart I [PDF, approx 23 Kb]). They are akin to the three foci or classes of evaluation, introduced by Rossi and Freemen (1989). They are crucial both at the planning stage and at the post program implementation evaluation or summative review[10]. In the planning stage it is important to use these questions to plan the data collection to be able to address these questions; and to anticipate the kinds of alternative interpretation of the results and the contentious issues which may arise to limit the effectiveness or efficiency of the program so as to counter them in

the design of the program and the evaluation.

In the summative or post hoc evaluation, it is important to have anticipated these issues and be prepared to report on them.

---

[10] If there is any doubt about the preparation of the proposal for the intervention, then Charts V and VI may be necessary before going to Chart I.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

2.2 THREE KEY ISSUES:
CAN X SUCCEED?
HAS X BEEN IMPLEMENTED PROPERLY?
WAS THE GRANT USED PROPERLY?

These key issues open up the door to good evaluation and clarity of purpose in injury prevention. The term 'evaluation' is applied to widely divergent types of work. As these produce widely differing types of information, and have very different requirements in terms of cost, time, expertise, it is essential to establish the purpose of an evaluation as early as possible in the process.

The heading of this subsection refers to three broad types of purpose which an evaluation may in general be expected to satisfy. Of course specific types of evaluation may address others more pertinent to the technicalities of the program or the auspice or funding agreement. Many introductions to evaluation consider these issues in various forms (see Owen, 1994). Rossi and Freeman (1989) astutely conceptualised these key issues as "foci" for classes or purposes of evaluation, which are quite suitable for the approaches commonly relevant to the evaluation of injury prevention and control programs. They pointed out that:

> "It is useful to distinguish between three major classes of evaluation research: (1) analysis related to the conceptualization and design of interventions; (2) monitoring of program implementation; and (3) assessment of program effectiveness and efficiency. Although it is not always possible to do so fully, the evaluation of social programs may need to include all three classes of activities. Evaluations that do so are termed comprehensive evaluations." (Rossi & Freeman, 1989, p. 44)

These questions form the backbone of Chart I (PDF, approx 23 Kb). From these foci the evaluation design can be directed to Chart II (PDF, approx 20 Kb) (based on McKillip, 1987, Figure 8.1, p. 102) for clarification of the needs and problems to be addressed. If there are uncertainties about suitable solutions for the injury risk or the difficulty in conceptualizing the evaluation in terms of the injury control intervention or prevention, then Chart III (PDF, approx 31 Kb) (based on McKillip, 1987, Figure 8.2, p. 103) is intended to provide a structured approach. As the evaluation attempts to draw on various data collection methods, Chart IV (PDF, approx 43 Kb) (developed from the GAO/PEMD 10.1.4, 1991) is a guide to selecting appropriate technique.

Thus the three key issues can be re-stated in terms of the three foci of evaluations of injury prevention and control interventions or programs:

1. *Can X succeed?* This is a focus on the conceptualization and design of the program or intervention X, usually for the purpose of evaluation of the **efficacy of X:**
   Rossi and Freeman (1989, p. 45) went on to identify further questions to be addressed in consideration of the design and conceptualization issues, which can be re-stated in terms of injury prevention and control as follow:
   a. Is an injury risk or intervention/prevention problem appropriately conceptualized?
   b. What are the extent of the injury risk or problem and the distribution of the target population?
   c. Is the risk identification or prevention program designed to meet its intended objectives?
   d. Is there a coherent rationale (e.g., causal model; or common social value) underlying it?
   e. Have chances of successful prevention program delivery been maximized?

f. What are the projected or existing costs and risks ?
g. What is the relationship between costs and risks vs benefits?

2. *Has X been implemented properly?* This is a question of monitoring of program implementation in terms of its inputs and processes.
   The issues to be addressed in consideration of this second focus in evaluation: on **implementation** issues, can be extended from Rossi and Freeman (1989, p. 46) in terms of injury prevention and control as follow:
   a. Is an injury control intervention or prevention program reaching the specified target population?
   b. Are the efforts expended on the injury control intervention or prevention program being conducted as they were specified in the agreement for the grant or the original program design?
3. *Was the grant used properly?* Is an assessment of the usefulness (utility) of the program X, usually in terms of **effectiveness and efficiency of X** with reference to its outputs and outcomes.
   Again the questions to be addressed in consideration of this third focus in evaluation: on **utility** issues, were well articulated by Rossi and Freeman (1989, p. 53) and can be stated in terms of injury prevention and control as follow:
   a. Is an injury control intervention or prevention program effective in achieving the intended goals?
   b. Can the results of the program be explained by some alternative process that does not include the prevention program or the identified risk factors or injury control procedures?
   c. What are the costs and efforts expended to deliver the injury control intervention or prevention program how do they compare to the benefits to program participants or target populations?
   d. Is the prevention program an efficient use of the granted funds or available resources, compared with alternative uses of these resources?

2.2.1 Can X Succeed?

"Can it succeed?" is used here as shorthand for a focus on evaluation projects that are seeking to establish and test new knowledge about an injury prevention method or intervention. The focus may be whether the method can achieve certain desired outcomes under optimal circumstances ('efficacy') or whether it does so in circumstances of more typical implementation ('effectiveness').

The key feature of these studies is that they are about causes and effects - i.e. does this intervention cause a reduction in that type of injury. Studies that address "cause and effect" questions in a serious manner are of central importance for injury prevention. They also tend to be difficult, expensive and time-consuming.

The minimum requirements for a "cause and effect" study are clear-cut definitions of the "cause" and the "effect", and sets of measurements of the amount of the "effect" with and without the "cause".

Typically, several factors conspire to make the reality of doing such studies complex. Definitions may not be very clear, or it may not be easy to achieve clear a distinction between groups with and without the "cause" or the "effect". Factors other than the "cause" being studied may influence the outcome. The condition being studied may be uncommon, difficult to detect, or slow to manifest itself.

In consideration of the possible answers to the question "Can X succeed?" Chart I provides some options for action, as follows:

A. if the answer is No (see option A on Chart I) then it is advisable to go back to clarify the problem or original need for the intervention or program (go to Chart II, based on McKillip, 1987, Figure 8.1 and refer to injury prevention models - see Gielen, 1992)
B. if the answer is "Don't know" (see option B in Chart II) then there is a reason for more research to clarify the solution or intervention chosen (go to Chart III based on McKillip, 1987, Figure 8.2; see also see Gielen, 1992)
C. if the answer if Yes or at least it works "Sufficiently" to be applied (see option C in Chart I) then it is appropriate to clarify the conceptualisation of the program and evaluation design before proceeding (go to the question 2 in Chart I).

Several major types of study design (with numerous minor variations) have been applied to "cause and effect" questions. The methods differ in their capacity to deal with the problems that afflict such studies.

Much attention is given to the potential of the methods to avoid producing misleading results because of the effects of factors other than the "cause" being studied. Randomised Controlled Trials (RCTs) are often regarded as providing the 'gold standard' in this respect (e.g., Boruch, 1997; Rossi & Freeman, 1989). A well designed and well

conducted RCT is capable of providing strong evidence on a question. Consistent results from a number of RCTs provide information about the questions which they address that is as trustworthy as any that is likely to be available.

In this design, people who are selected to be the target 'cases' (i.e. exposed to some factor, such as a preventive intervention) and 'controls' (i.e. a comparison group) are chosen at random from the group to be studied (see Boruch, 1997; Rossi & Freeman, 1989). Preferably no one knows who has been selected for each group until the end of the study, though this is only possible for some types of intervention. The intervention is then undertaken, and things that it is expected to affect are measured. Other factors are usually measured as well, to help ensure that they do not lead to false conclusions. A key advantage of the RCT design is that the random allocation of people as cases and controls should ensure that unknown factors which might lead to false conclusions are distributed fairly equally between the two groups. This greatly (and to a predictable extent) reduces the chance that such factors will distort study findings to a serious degree.

But RCTs are not an guarantee of valid evaluation, nor a cure-all for problems of interpretation (Boruch, 1997; Campbell & Stanley, 1963; Campbell & Boruch, 1975; Rossi & Freeman, 1989). Indeed there is a major ethical concern (see American Evaluation Association, 1990; Boruch, 1997; Joint Committee on Standards, 1994; Sharp, 1994) with prospective case control or comparison group studies in the selection of those cases to be treated with the desired preventative approach, because those cases who are "unlucky" or not "randomly" selected, also have as much right to the prevention of injury as anyone else (Dannenberg, and Fowler, in press, 1998; Hollister & Hill, 1995). In retrospective studies (e.g. following up victims or using old case histories) or large sample content analyses, there is also the cost and the uncertainty about sample size needed to reach the degree of rigour for the randomization of selection of cases to be beneficial in the interpretation of causal linkages (Boruch, 1997, Covey, 1982; Dannenberg, and Fowler, in press, 1998; Hollister & Hill, 1995).

Other factors which mitigate the decisions are *scope* of question, timeliness, audience expectations (see Rossi & Freeman, 1989; Patton, 1997)

2.2.2 Has X been implemented properly?

This is a question of management, which is often subject to *auditing*[11] as well as evaluation (see Arens, Loebbecke, Best & Shailer, 1990; Anthony, 1988; Anthony & Young, 1988; Atkinson, 1993; Ballard & Macartney, 1995). In the context of evaluation, the answers to implementation issues often depend on the degree of monitoring of the ongoing performance of the program.

Monitoring is the regular gathering of data about the performance and operations of the program. As distinct from evaluation, which seeks to make judgements about the program, monitoring is a basic measuring and surveillance process (see Rossi & Freeman, 1989; Sharp, in press 1998).

Among the many reasons for monitoring programs are the following:

- proper administration and management of programs require that managers, in their routine activities, conduct their activities as efficiently as possible, vigilance in detecting errors and monitoring waste as well as performance are essential in management of all programs (see Anthony & Young, 1988; Love, 1991; Rossi & Freeman, 1989)
- accountability to the auspicing (funding) agencies or program sponsors and stakeholders, require evidence that the program was actually delivered; and financial probity demands that managers account for what was paid and what was actually undertaken. Monitoring provides vital feedback data and documentation of the operations and evaluation uses this data as information for management on the efficiency and effectiveness of the operation of the program (see Donovan & Jackson, 1991; Stufflebeam, 1971)
- in order to be able to interpret the validity of the intervention and its usefulness or impact or outcome a program manager or Evaluator has to first demonstrate that it took place in the planned or appropriate manner and for the target participants. This is confirmation of the necessary link between planning and action.

2.2.3 Was the grant used properly? Or was the intervention useful?

This evaluation focus is probably the most often considered. If an evaluation of injury prevention programs does nothing else it is likely to be concerned with utility questions, especially as a criterion of evaluation of a program or

intervention (e.g., Elvik, 1997; Selby, 1994; Zwerling, 1997) and as a criterion for meta-evaluation, i.e. how useful was the evaluation for decision-makers (e.g., Patton, 1997; Hollister & Hill, 1995; Joint Committee on Standards, 1994).

In addressing these evaluation foci, the design of the program needs to incorporate the design of the data collection and follow the logic of the inference in program logic terms. Chart IV outlines some of the decisions and processes which are involved in the program logic of the choice of evaluation design and data collection method. This chart, like most of the evaluation and injury prevention research fields, makes assumptions about the hierarchy of validity or strength of inference of the evaluation design. The next section introduces some of the issues involved in those design processes and methods.

NOTES on CHART VI

Chart VI (PDF, approx 32 Kb) is introduced here on the prospect that there may be considerable uncertainty in the preliminary decisions about how or whether to select an intervention program proposal. To address the needs of program Administrators or Managers who may have to become evaluation practitioners, the prior issues of program selection are included in Chart VI.

The fields of medical research have benefited greatly from a systematic approach to documenting and communicating reviews of key findings, through the Cochrane Collaboration (see Clunie, Ludbrook & Faris, 1995; Sligay & Jewell, 1994). One of their many contributions has been to provide a clear standard by which to prepare and judge literature reviews, as outlined in Table 4. Such approaches can be very useful in evaluating a program proposal.

Also Sackett, Haynes & Tugwell (1985, p288 & 23 & 230) have identified a useful flow chart for reviewing clinical articles for epidemiological reviews, and basic pictorial and verbal explanations of the various quasi-experimental and experimental designs (eg. RCT).

TABLE 4

THE COCHRANE COLLABORATION
What is Systematic Review?

1. "State objectives of the review, and outline eligibility criteria
2. Search for studies that seem to meet eligibility criteria
3. Tabulate characteristics of each study identified and assess its methodological quality
4. Apply eligibility criteria, and justify any exclusions
5. Assemble the most complete: dataset feasible, with involvement of investigators, if possible
6. Analyse results of eligible studies. Use statistical synthesis of data (meta-analysis), if appropriate and possible
7. Perform sensitivity analyses, if appropriate and possible
8. Prepare a structured report of the review, stating aims, describing materials and methods, and reporting results"

Source:
Cochrane Collaboration (1996) "What is a Systematic Review?" Cochrane Collaboration September 23. 1996
http://hiru.mcmaster.ca/cochrane/

Another set of criteria which are useful in determining the degree of credibility of a proposal would be the extent of compliance with Hill's Criteria of Causation. (see Table 5).

TABLE 5

HILL'S CRITERIA OF CAUSATION

The first complete statement of the epidemiologic criteria of causal association is attributed to the British medical statistician Austin Bradford Hill (1897-1991), although others enunciated several of them. The criteria of a causal association of a factor and a disease are:

1. *Consistency:* The association is consistent when results are replicated in studies in different settings using different method.
2. *Strength:* this defined by the size of the risk as measured by appropriate statistical tests.

3. *Specificity:* this is established when a single putative cause produces a specific effect.
4. *Dose-response relationship*: An increasing level of exposure (in amount/or time) increases the risk.
5. *Temporal relationship:* Exposure always precedes the outcome: This is the only absolutely essential criteria.
6. *Biological plausibility:* The association agrees with currently accepted understanding of pathological processes. This criterion should be applied with caution. As Sherlock Holmes remarked to Dr. Watson, "When you have eliminated the impossible, whatever remains, however improbable, must be the truth".
7. *Coherence:* The association should be compatible with existing theory and knowledge.
8. *Experiment:* The condition can be altered (eg. , prevented or ameliorated) by an appropriate experimental regimen.

Source:
Last, J.M. (ed.). 1995. *A Dictionary of Epidemiology*. Oxford University Press, New York, p. 77

These criteria should strengthen the usefulness of Chart VI as a guide to evaluation of a proposal.

---

[11] Auditing is a field unto itself and too complex and intense to be covered here (see e.g., Arens et al. 1990; Sawyer & Scheiner, 1996). However, there are significant relationships and convergence between internal audit and program evaluation (see Hudson & McRoberts, 1984; Love, 1991; Pollitt & Summa, 1996; Ryan, 1993; Schwandt, & Halpern, 1988), especially in relation to program implementation issues.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

2.3 HIERARCHY OF EVALUATION DESIGNS

Since the early days of evaluation, when it was quite often called "evaluative research" (Hyman, Wright & Hopkins, 1962; Suchman, 1967) or "evaluation research" (Weiss, 1972; Rossi & Freeman, 1989), there has been a strong deference to the traditional scientific research paradigm (see Hughes, 1980; Kaplan, 1964; Kuhn, 1970; Lakatos, 1970). For example, the influence of the classic paper in educational research by Campbell and Stanley (1963) has permeated not only other social sciences but also dominated evaluation design and meta-evaluation for decades (e.g. Hudson & McRoberts, 1984; Isaac & Michael, 1995; Rossi & Freeman, 1989). The basic thrust of this paradigm is that evaluation is intended to assist in interpreting the effectiveness of a program or intervention. In other words: did the program, or intervention, actually produce (cause) the results that were observed? Also it was important to ask whether or not the results were what was intended? In order to interpret the results of an evaluation in this dominant paradigm, the design of the sequencing of events of the program, as well as the methods and opportunities for data collection, must be such that the logic of the inference of causal linkages between what was intended as an intervention and the results obtained, can be sustained against a set of standard "plausible rival hypotheses" (Campbell & Stanley, 1963). They advised that variations of the "controlled" experimental research design could be as useful in this inference process, provided the researcher (or Evaluator) knew the degree of control and the extent of invalidity from plausible rival hypotheses which could intrude in the quasi-experimental design. The stronger the control which is able to be exerted over the sequencing of events of the program, and the ability to manipulate the evaluation methods, is thought to increase the strength of the logic of the inference of causal linkages (Davis, 1985; Kuhn, 1970; Lakatos, 1970; Platt, 1964). The scientific argument proposes that the stronger the inference of relationship between program intervention and the results obtained, the greater the anticipated accountability of the program for its resources; also the greater the predicability of future results, and future estimation of resource requirements (see Campbell, 1969; Davis, 1985).

This paradigm has led to many attempts at ranking the strength of the designs of programs and evaluation methods to enhance the strength of the inference from the results. For example in the health promotion field, Green (1979) has suggested the following "hierarchy of evaluation designs", moving from 1 to 6 in terms of increasing control of the variables and the resulting strength of inference:

1. record-keeping (historical approach)
2. stock take (inventory approach)
3. comparative approach
4. controlled comparison (quasi-experimental design)
5. controlled experimental design
6. double-blind randomised controlled trial (RCT) design

Rather than attempt a comprehensive analysis of evaluation, quasi-experimental and experimental research designs which have been provided by many (see Boruch, 1997; Campbell, & Stanley, 1963; Cook & Campbell, 1979; Green, 1979; Lewis-Beck, 1993; Sackett, et al, 1985), it is more relevant here to address the issue of how to classify the studies reviewed. In his useful review of the design and use of literature reviews Cooper (1989; Cooper & Hedges, 1994) identifies three approaches to the categorisation of research (and evaluation) methods, viz:

- **"threats-to-validity approach":** as identified by Campbell & Stanley (1963) the internal and external validity of experimental and quasi-experimental research (and evaluation) designs are susceptible to various threats from standard "plausible rival hypotheses"; Cook and Campbell (1979) identified 33 such threats to validity which are beyond the scope of analysis of a review such as this;
- **"methods-description approach":** a thorough description of the characteristics of the method identified by the original researchers; this involves less interpretation but more detailed description;
- **"mixed-criteria approach":** as is suggested, this more eclectic approach is based on a comprehensive description of the methods reviewed, but allows a diagnosis of threats-to-validity as they emerge.

This third option, which Cooper (1989) called "optimal", will be adopted for the purposes of the current literature review.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

2.4 EXAMPLES OF INJURY PREVENTION REVIEWS

An important review of childhood injury (Rivara, Beahler, Patterson, Thompson, & Zavitkovsky, 1997) has been established as a link with the Cochrane Collaboration, by the Harborview Injury Prevention Research Centre, University of Washington (see http://weber.u.washington.edu/~hiprc/childinjury/).
In addressing the *hierarchy of evidence* issue, they listed studies according to the following designs (ascending in order of strength of inference):

1. *Ecologic study* (the intervention or risk is exposed to an entire population or group) although there is no control of the exposure some non-intervention comparison groups may suffice for comparison, the timing and the type of group provide the following bases for comparison:
    A. Ecological time (or time-tend) study: i.e., a before and after comparison of study.
    B. Ecological group study (comparing two or more groups concurrently)
    C. Ecological mixed study (between and within group comparisons over time)
2. *Case-control study*, where the "cases" are people *with the outcome of interest*, versus "controls" who are people thought to be comparable to the cases, but *without* the outcome of interest (e.g. uninjured workers from the same workplace as injured "cases"). The two groups are analysed for differences in levels or types of exposure to possible risk factors or protective factors.
3. *Cohort study* (comparisons of different interventions or exposures to risk factors, for samples of the same cohort, or *panel* of respondents)
4. *Controlled* trials that were *not* randomized (convenience samples, or whole cohorts)
5. *Randomized controlled trials* (including some community trials)

However, Rivara et al (1997) decided not to include the following study types of studies:

- Case series (anecdotal or uncontrolled cases)
- Laboratory studies (non-human or inanimate objects were the subjects)

It is a pity they discarded case studies in this way. Case studies can be useful in evaluations (see below section 2.4.4) provided that various conditions like Hill's criteria are met (see e.g., Kazi, 1996; Sechrest, et al., 1996; Rossi & Freeman, 1989).

While Rivara et al (1997) provide a model of how to categorise such studies, the present review is more diverse in its search and is intended to be primarily heuristic as well as to account for rigour. Thus the "mixed-criteria approach" (Cooper, 1989) is most appropriate here. Certainly it is quite often recommended in the evaluation literature (see Patton, 1990; Scriven, 1991), and recently promoted in the health promotion literature (see Macdonald, 1996).

2.4.1 Randomized Controlled Trial

The chief concern for quantitative studies of causes and effects is that an observed effect may be due to a factor or factors other than the one of primary interest. Several study designs incorporate comparison groups to reduce the chance of drawing false conclusions because of this type of problem. The study design capable of providing the most rigorous defence against this is the Randomised Control Trial (RCT), in which subjects are allocated at

random to a group to be exposed to the factor being studied (cases) or to a control group. If possible, subjects and investigators are "kept in the dark" about whether each subject is a "case" or a "control" until the end of the experiment (i.e. a "double-blind" trial). RCT's have important strengths, and substantial limitations. (see Campbell & Boruch, 1975) They tend to be expensive and time-consuming (it is unusual to be able to use existing data). Many questions cannot be studied by this method for ethical and logistic reasons. For Example, The best evidence on the effectiveness of bicycle helmets comes from "weaker" study designs than RCT's. A decision to conduct an RCT on this subject would have to confront the ethics and practicality of establishing, by random allocation, a group of "wearers", and a group of "non-wearers", and following their injury experience for a long period (probably years).

Typical examples of this approach are the focus of the Cochrane Collaboration (see e.g., Clunie, Ludbrook & Faris, 1995; Silagy & Jewell, 1994) focusing on primary health care and general medical practice.

In injury prevention research and evaluation it is not always as applicable or less systematically applied in various fields. For example, Zwerling et al (1997, p. 164) found that in the occupational injury intervention studies they reviewed, that: "randomized controlled trials are rare and also ... that the quasi-experimental studies in the literature often use the weakest designs"

However, as Dannenberg, and Fowler (in press, 1998) point out the truly double-blind, randomized control trial design is less frequently used in injury prevention evaluation because of the ethical, financial and logistical constraints. Indeed, these are common problems among evaluations of other community based social interventions prevention (e.g., Covey, 1982; Hollister & Hill, 1995).

2.4.2 Case-Control

Here the "case" person (or group) who has suffered an injury is compared with their ilk who did not (the "controls") in terms of exposure to one or more suspected "risk factors" for the injury. Armenian and Lilienfeld, (1994) provide an excellent historical review of the use of case-control methods. This is one of the oldest and most frequently used designs (see Solomon 1949). Because it only relies on the existing selection of the case and the control, which are not randomly allocated to groups, there are a number of plausible rival intervening variables, including selection bias, which could explain the findings of such a study. Despite its perceived limited *scientific* rigour, it is still a useful evaluation design, because it enables some basic comparison to establish change and relative effectiveness (see Armenian & Lilienfeld, 1994; Rossi & Freeman, 1989).

2.4.3 Cohorts or Panel Surveys

The group of people enrolled in the study have various levels of exposure to a suspected risk factor of interest (there may only be two levels: exposed or not exposed). They are followed-up for some time, and the occurrence of injury is recorded. Analysis is designed to look for association between exposure and risk of injury.

A series of measurements or surveys of a group of similar individuals at risk of injury, providing comparisons over time, before and after such risks might occur, can be a useful technique for better understanding the impact of risk factors, the needs of those at risk, and for pilot testing models or interventions. It can be a preliminary basis for testing the efficacy of the injury prevention (e.g. Berry, Gilmore & Geller, 1994) In some literature this is called a panel survey (see GAO/PEMD 10.1.4 , 1991; Rossi) Berry, Gilmore and Geller (1994) provide two examples of time series analysis of longitudinal case studies of automobile seat belt use.

2.4.4 Individual-subject or single case

Single case study (see Fisher, 1988; Hamel et al, 1993; Kazi, 1996; Sechrest, et al., 1996; Rossi & Freeman, 1989; Smith & Everly, 1988; Yin, 1994) is often associated with the Naturalistic inquiry and the qualitative approach to evaluation (Lincoln & Guba, 1985; Guba & Lincoln, 1990; Patton, 1990). Sechrest, Stewart, Stickle, & Sidani (1996) and Kratchowill (1979) argue and demonstrate that the case study can have a high degree of validity and interpretive power, given that they meet certain criteria of evidence.

Despite its complicated subjectivity and all the limitations of this method, it has one important advantage over other evaluation designs, viz. its ability to capture the imagination of the media, the public and decision makers by the perception of the human interest, and the personal impact of the injury or the injury prevention or control

intervention. This is the stuff of the front-page photo or anecdotal opening remarks for the executive summary or preface which can grab the attention of the decision-makers and bring home the need for the injury prevention intervention. Of course the findings could go for or against the plans of the injury prevention or of the evaluation. That is the risk of this kind of study, viz the inability to control the variables and the interpretation of the results, as well as the ease with which it can be dismissed by critics.

The ethical issues associated with this kind of study are probably more significant than the usual evaluation, because of the exposure of the participant to such attention (see Joint Committee on Standards, 1994; Sharp, 1994). Also there are difficulties in mixing case studies of at-risk participants with the more management or efficiency oriented focus of some evaluations (e.g. See Gray, Marshall, & Morris, 1997).

Sechrest et al. (1996, p 4 - 4) point out that:

> "Remarkably little research has been done on the effectiveness or persuasiveness of case studies. In fact, we know of only one such study, a dissertation carried out by David Ametrano at the University of Michigan. That is, we think, a highly unfortunate gap in our knowledge of a widely used and potentially important methodology. Obviously more research needed if we are to be able to write dependable specifications for producing persuasive case studies."

Based on their research, they argue that there are at least three elements which can improve the persuasiveness of case studies, as demonstrated in "classic" fundamental cases.

> "First, it is true that cases that have come to be widely known and accepted as generally plausible are interesting in their own right. ...
>
> A second characteristic of classic case studies, most of which are in one way or another meant to be probative, is that they either include or are embedded in a context that represents a strong theoretical basis for inference. ...
>
> A third characteristic of classic case studies is that they centred around major, sometimes cataclysmic, events with very large and obvious consequences. ..."
> (Sechrest et al. , 1996, p 4-4)

Kazi (1996) also demonstrates that a combination of qualitative and quantitative data collection techniques can assist the Evaluator to make a more persuasive interpretation of the case study.

FIGURE 7: Overview For Planners Linking Program & Evaluation Plans for Injury Prevention Programs – A Guide to Charts I to VI

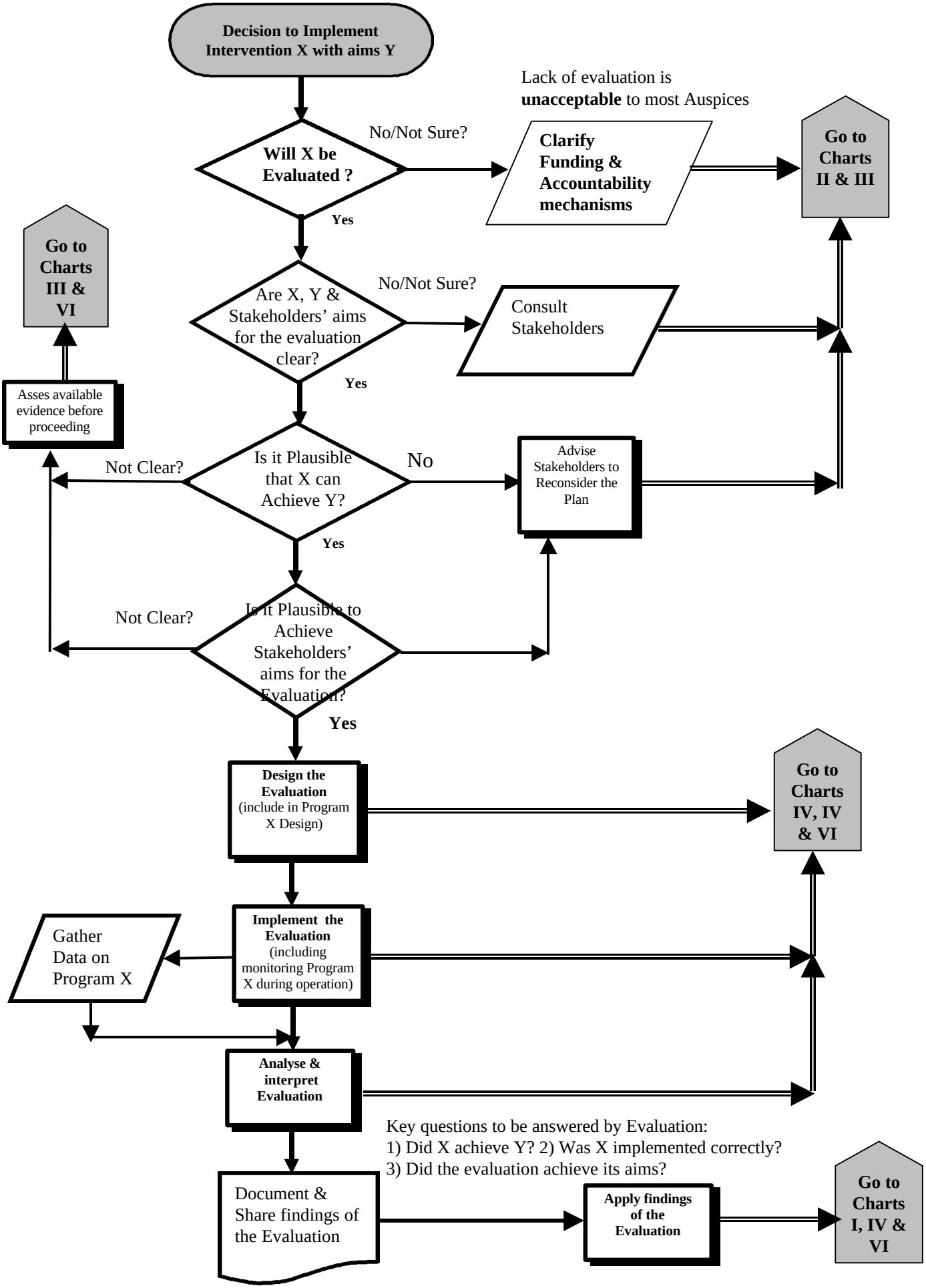


Chart I: For Planners Linking Program & Evaluation Plans to Evaluation Foci for Injury Prevention Programs

Based on Rossi & Freeman (1989, pp 44-52)

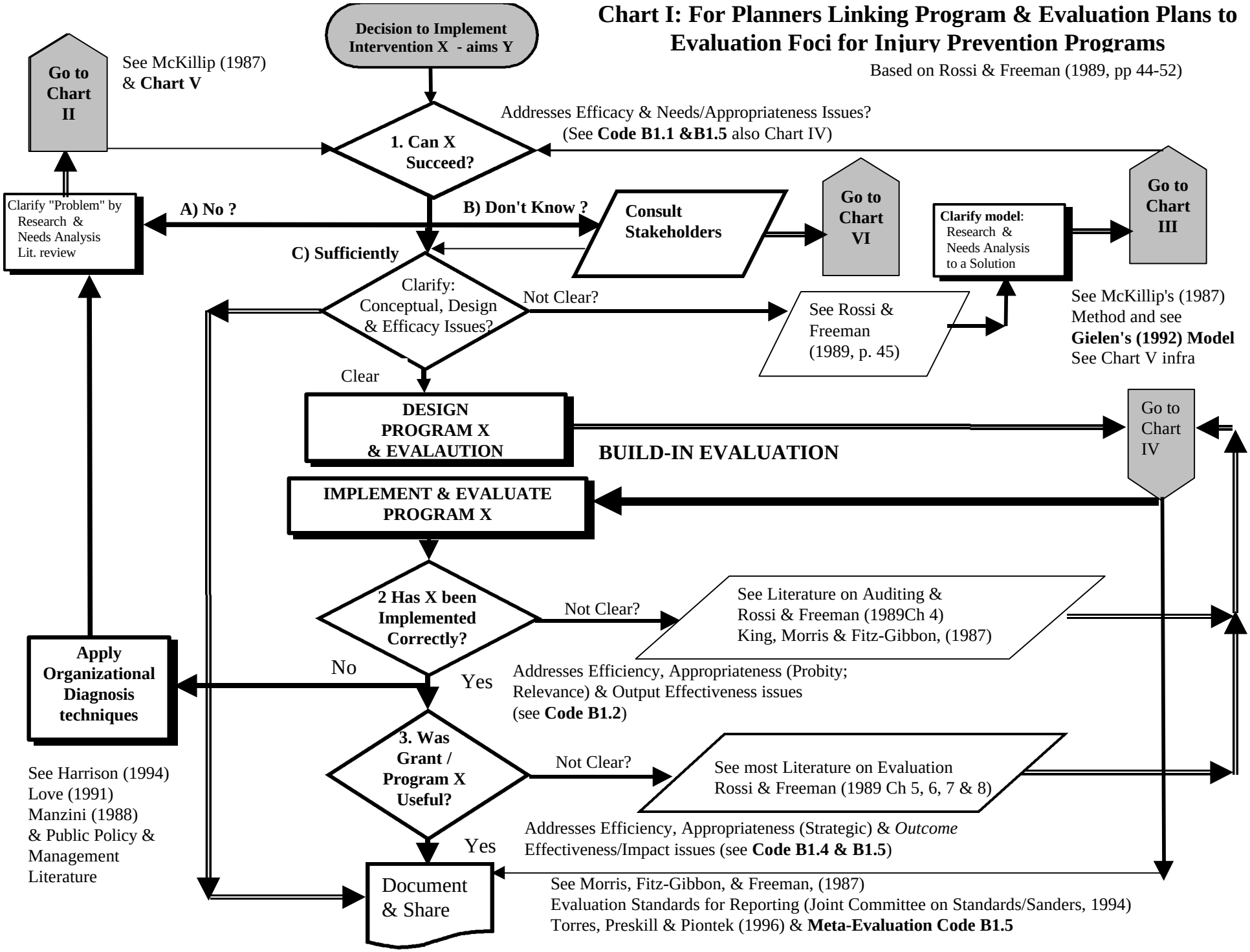


Chart II: Identifying Problems in Planning & Evaluation of Injury Prevention Programs

Based on Mc Killip (1987, p. 102)

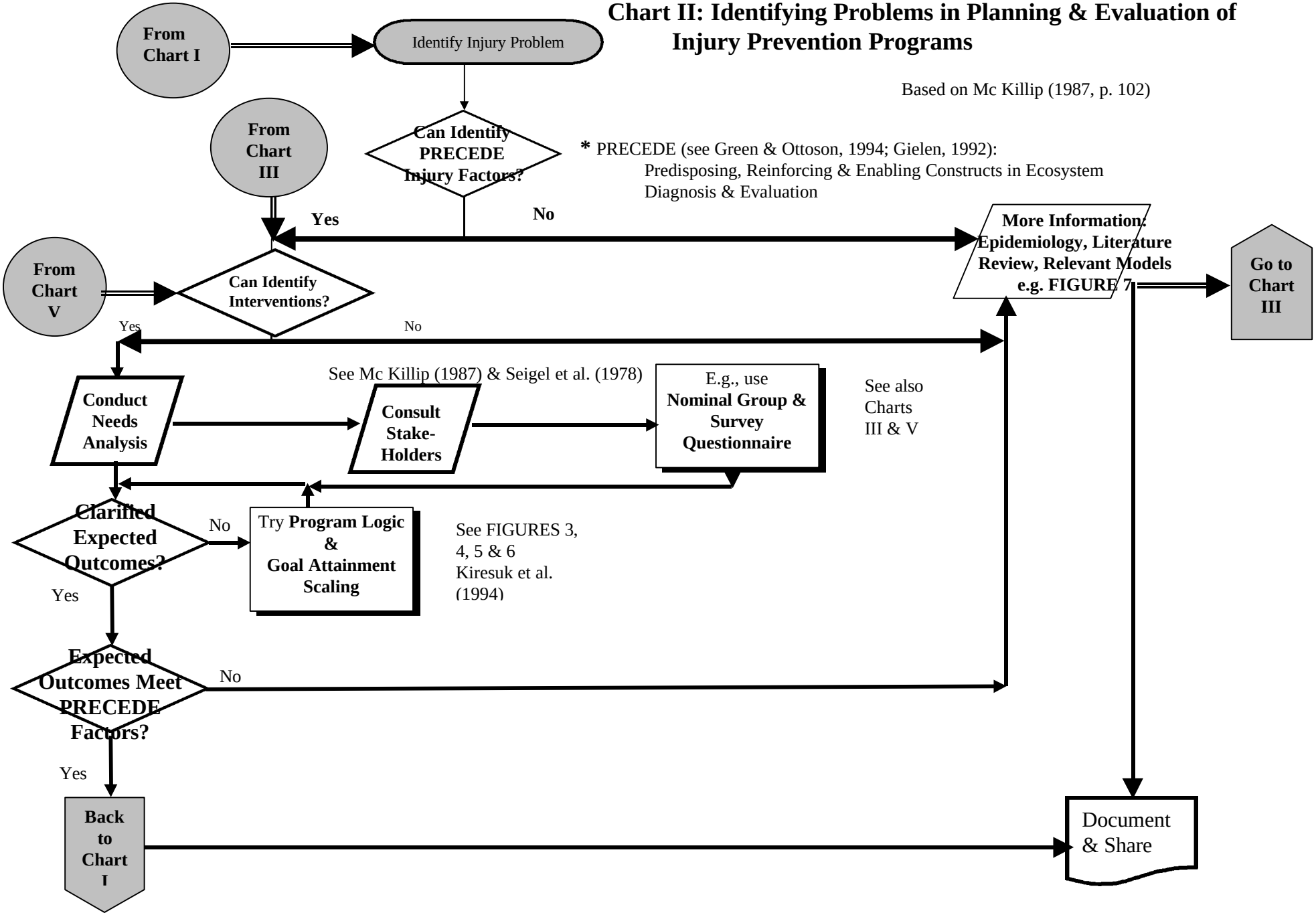


Chart III: Identifying Solutions in Planning & Evaluation of Injury Prevention Programs

Based on Mc Killip (1987, p. 102)

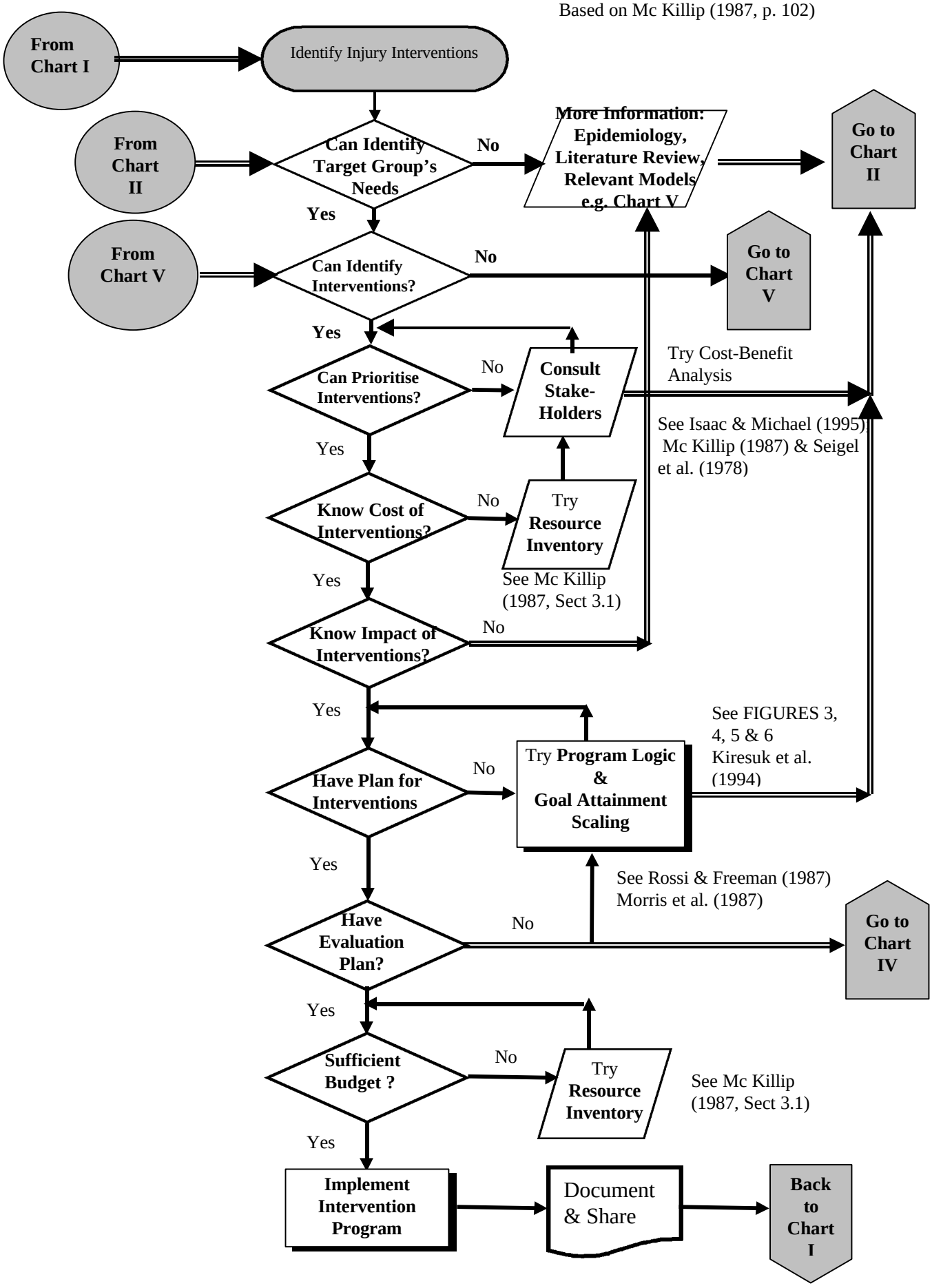


Chart IV: Linking Evaluation Design to the Evaluation Questions

Based on GAO/PEMD 10.1.4 (1991, pp 68 -69)

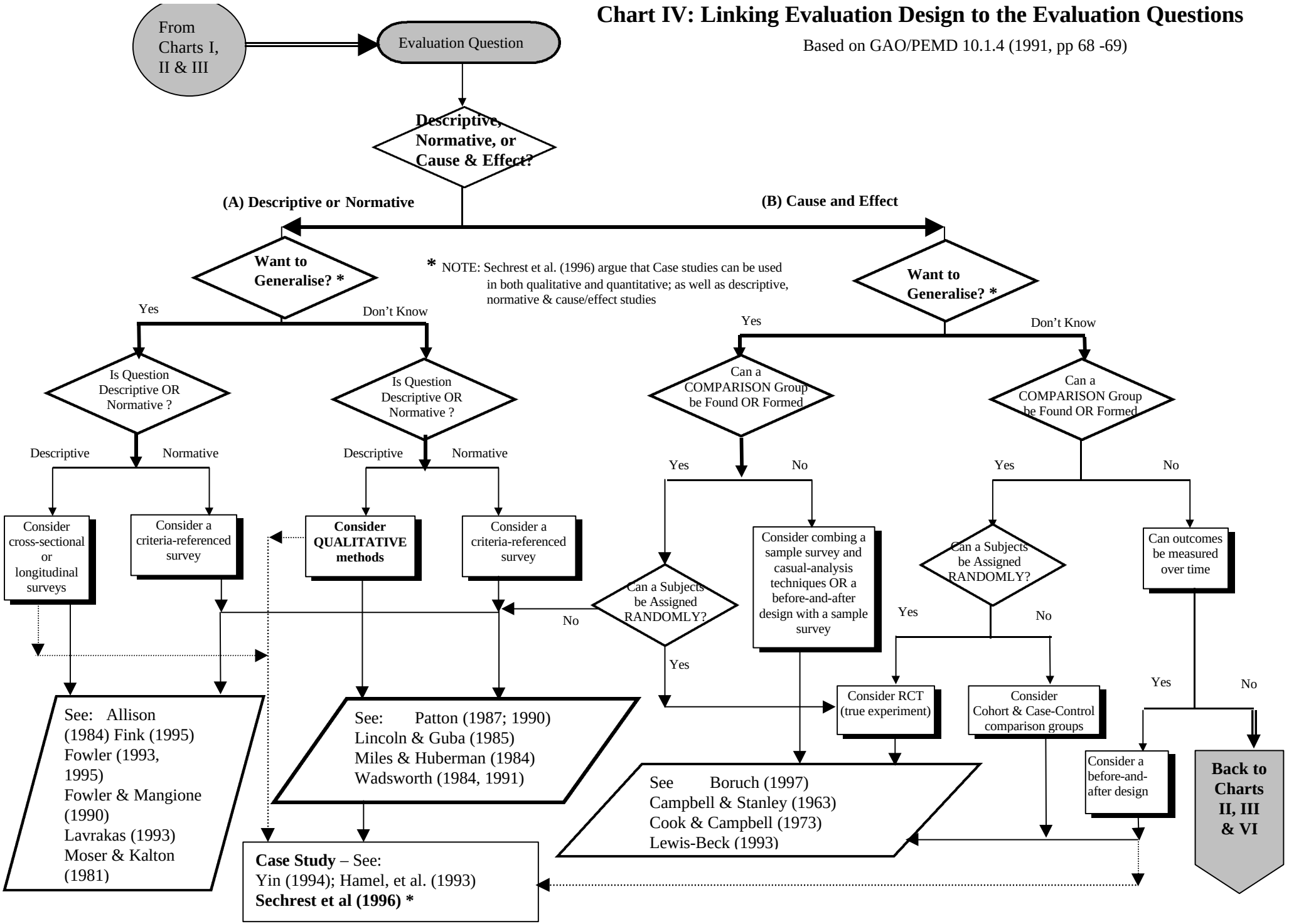


Chart V: Application of Gielen's (1992) Integrative Planning Framework for Injury Prevention Programs

To An Example of Childhood Motor Accident Vehicle Accident Prevention
(Based on Gielen, 1992, pp 208 & 210; Green & Ottoson, 1994)

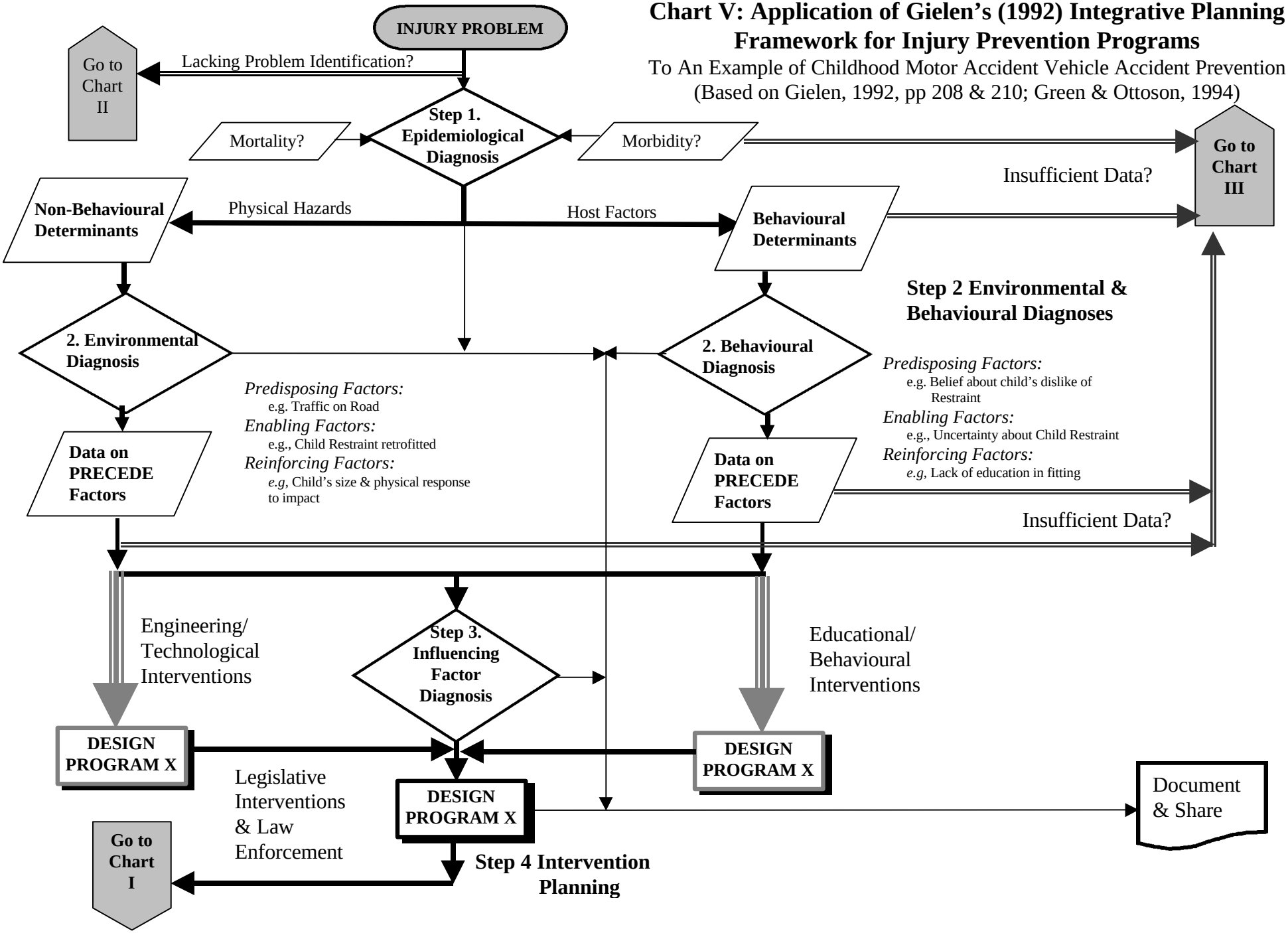
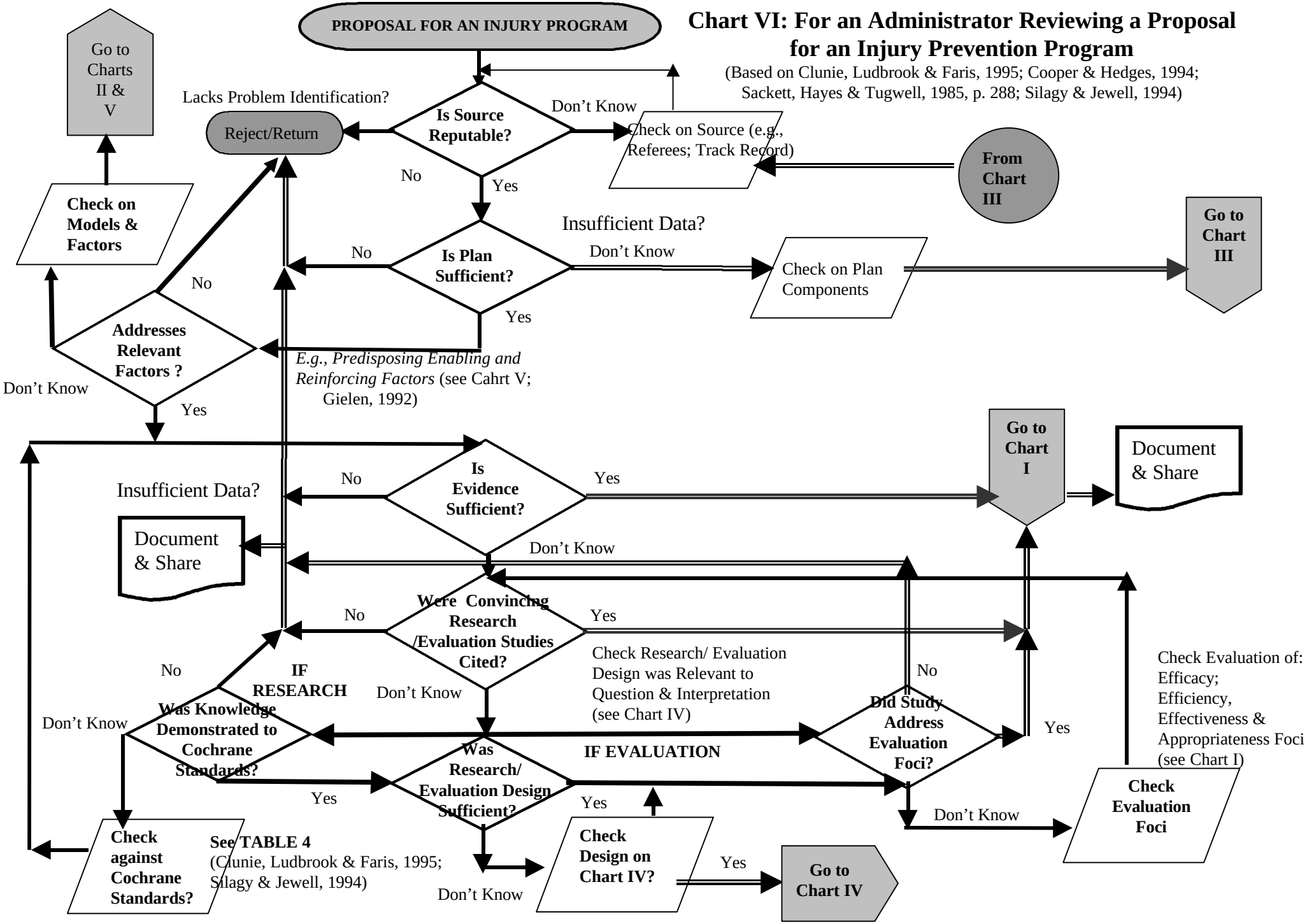


Chart VI: For an Administrator Reviewing a Proposal for an Injury Prevention Program

(Based on Clunie, Ludbrook & Faris, 1995; Cooper & Hedges, 1994; Sackett, Hayes & Tugwell, 1985, p. 288; Silagy & Jewell, 1994)



FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

4.1 Aim

The main aim of this project is to address the needs of potential evaluation practitioners in the fields of injury prevention and control. To do this we illustrate the range of evaluation models and techniques, and encourage opportunities for their application, by finding examples of evaluation literature, relevant to injury prevention. It is not feasible to attempt a comprehensive review of all the literature available. So we have tried to at least gather a few examples of each of the main models in Table 3.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

4.2 Search Methods Used to Find Articles

Relevant articles were identified using several sources. Electronic database searches were conducted via the internet and CDROM (Winspirs) databases. The following databases were searched over the internet: Uncover, Academic ASAP, and Medline. The following databases were searched using CDROM (Winspirs) databases: Medline, Sociofile, Psyclit, Eric, CINAHL, and ABI Inform.

Some of the main keywords used in the data base searches were:

- evaluation of injury prevention;
- evaluation designs;
- evaluation of community based interventions;
- evaluation of outcomes;
- evaluation and accidents;
- burns prevention;
- occupational health and safety evaluation;
- problems with evaluations;
- evaluation and interventions;
- evaluation and analytical epidemiology;
- environmental evaluation;
- biomechanical evaluation;
- evaluation of health promotion;
- evaluation policy;
- evaluation process;
- quantitative evaluation;
- qualitative evaluation;
- evaluation of training and injury prevention programs;
- quasi experimental designs;
- and specific injury categories, such as fractured neck of femur.

In addition to the above database searches, references to articles were obtained from contact with people via listservers, over the internet, telephone interviews with experts in the area of injury prevention, searches of bibliographies of articles obtained, and some hand searching of recent editions of relevant journals such as *Evaluation Review; New Directions for Evaluation*; and *Accident Analysis and Prevention*.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

4.3 Review of Articles

Articles were reviewed to determine their suitability for the bibliography. Articles were included if they were considered to be a "good example" of an application of an appropriate evaluation method, or if they presented a discussion of appropriate evaluation methods, or if they were comprehensive meta-analyses of evaluations which had been carried out. Although many of the papers deal specifically with the area of injury prevention, the literature review was not restricted to this area only, as there are relevant examples of appropriate evaluation methods from other fields. Papers which were included in the bibliography were classified according to the four dimensions which appear in Section 4.2 below. If a paper has not been included in the bibliography it does not mean that it was an inappropriate article, as it could be that there are a significant number of papers not gathered by the review process. However, those papers referred to here are a sufficient basis on which to address the terms of reference of the brief provided.

The articles were reviewed by two reviewers with backgrounds in the social sciences and program evaluation, and referred to the Client (NISU) representative for a third opinion or advice on technicalities of the injury prevention literature, where necessary.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

4.4 Key Categories of Papers to be Classified

Classifications of the articles cover the two broad categories of

A. concerning the injury program/project that is being evaluated (including the subject or topic of investigation and the type of program)
B. concerning the evaluation (its criteria and methodology)

These broad categories are explicated below, with each of **A** and **B** being broken down into two sub categories. Note that each element under the sub-headings is given a specific code number, which will be used to classify the papers chosen for the bibliography.

TABLES IDENTIFYING THE LITERATURE REVIEW CATEGORY CODES

Category A1: Subject/Topic Matter

Code No.	Subject/Topic Matter
A1.1	Occupational
A1.2	Drugs/Alcohol
A1.3	Falls Prevention
A1.4	Childhood Injury
A1.5	Scolds/Burns
A1.6	Agricultural Injuries
A1.7	Road Accidents
A1.8	Explosives/Fireworks
A1.9	Mixed
A1.10	Non-Injury Related Area

Category A2: Main Preventive Approach Used

Code No.	Preventive Approach
A2.1	Environmental
A2.2	Biomechanical
A2.3	Health Promotion and Education
A2.4	Community Development
A2.5	Policy
A2.6	Legislation
A2.7	Mixed
A2.8	Non Injury Related Field

Category B1: Evaluation Foci (Criteria or Objectives)

Code No.	Criterion	Explanation of Category

B1.1	Efficacy	Evaluation is to determine whether an intervention or a component of an intervention has its intended effects under "ideal" circumstances
B1.2	Effectiveness of *Implementation:* Inputs vs Processes or Outputs	Monitoring (Auditing) & Evaluation is to determine whether study/program *inputs* cause changes in specified *processes* or *outputs* which they are intended to influence
B1.3	Effectiveness of *Outcome*: Inputs or Process vs. Outcomes	Evaluation is to determine whether the study/program *inputs* or *processes*(interventions) cause changes in outcomes (intended or unintended)
B1.4	Efficiency	Evaluation is to determine whether an intervention or program (process) which has shown to be effective is better than others in terms of the amount of *input* required to achieve a specified level of *process change, output* or *outcome*
B1.5	Appropriateness	Evaluation is to assess the match between program *objectives* or intended outcomes and the *needs* and expectations of the people to whom they are intended to apply, or to other stakeholders.
B1.6	Multiple (Comprehensive)	Evaluation foci may include more than two of the other criteria
B1.7	Meta-Evaluation	The audit or evaluation of the evaluation, to determine whether the quality of evaluation meets identified standards

Category B2: Evaluation Orientation(s)

These categories are not mutually exclusive and reflect the general literature on evaluation, as identified by various reviewers (see also Table 3 on models of evaluation).

Code No.	Orientation(s)	Explanation of category
B2.1	Measurement	Quantitative monitoring; no quantitative target, but could develop a normative database
B2.2	Objectives	Quantitative monitoring with respect to a stated target value
B2.3	Needs	Evaluation in terms of an implicit goal of eliminating a particular "need" or Satisfying consumers' expectations
B2.4	Management	Evaluation focused on providing what is needed to enable certain decisions to be made
B2.5	Causality	Research and Evaluation aimed at testing causal hypotheses about an intervention and its effects
B2.6	Expert Opinion	Evaluation based on determination of the views of an expert or panel of experts, or consensus/conference of stakeholders
B2.7	Self-Evaluation	Evaluation of a program by its management/staff for their own purposes and program improvement
B2.8	Responsiveness	Evaluation intended to clarify program values and processes, and stakeholder views
B2.9	Utilisation	Evaluation focusing on identifying intended uses of the evaluation by stakeholders and producing an evaluation that is useful and to be used
B2.10	Meta - Analysis	Evaluation which systematically searches out and combines the results of studies that have addressed a similar research question, in order to distill the basic common findings
B2.11	Ecclectic/Comprehensive	The paper may combine the various orientations

Note that some of the articles selected for the bibliography may cover more than one sub-heading in each of the four categories mentioned above (e.g. the evaluation objectives may be both efficacy and appropriateness).

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

REFERENCES CITED IN THE INTRODUCTION

American Evaluation Association (1990) "Standards and Ethics Committee". *Evaluation Practice*, Vol. 11, No.1, pp.96-97.

American National Standards Institute (1989): *American National Standard Method of Recording Basic Facts Relating to the Nature and Occurrence of Work Injuries Z 1 6.2.* New York, NY: ANSI.

Attkisson, C.C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds) *Evaluation of Human Service Programs.* N. Y., Academic, 1978

Australian Council of Social Services (1978) "What is evaluation?" *Australian Social Welfare,* March, pp. 16-18.

Australian Youth Foundation & Sharp, C.A. (1996) *Do-it-yourself Evaluation Manual: An integrated approach to project management and evaluation.* Sydney. Australian Youth Foundation.

Auslander, G.K. (1996) "Outcome evaluation in host organisations: A research agenda" *Administration in Social Work*, vol. 20 (2), pp. 15 - 27.

Baume, P. (1991) "Getting the policy right: Recovering the golden rules of evaluation" In J. Uhr (Ed.) *Program Evaluation.* Federalism Research Centre, The Australian National University, Canberra, pp. 35 - 39.

Bartos S. (1995) "Current developments in performance information" *Australian Journal of Public Administration* vol. 54 (September): pp. 386 - 392

Boruch, R.F. (1997). *Randomized experiments for planning and evaluation: a practical guide*, Thousand Oaks, Sage.

Brinkerhoff, R.O. & Dressler, D.E. (1990) *Productivity measurement: a guide for managers and evaluators*. Newbury Park, Sage.

Bruell, J.D. (1994) "How 'The Government Perfomance and Reults Act' Borrows from the Experience of OECD Countries" *New Ideas, Better Government - The Fulbright Symposium on Public Sector Reform*. Centre for Australian Public Sector Management, Griffith University, Brisbane (23 - 24 June, 1994).

Burdorf, A. Sorock, G., Herrick, & Courtney, T.K. (1997) "Advancing epidemiological studies of occupational injury - Approaches and future directions" *American Journal of Industrial Medicine*, vol. 32, pp. 180 - 183.

Camp, R. (1989) *Benchmarking: The search for Industry Best Practices that will lead to Superior Performance.* Milwaukee, Wisconsin, Quality Press.

Campbell, D. & Stanely, J.C. (1963) *Experimental and Quasi-experimental Designs for Research.* Chicago, Rand-McNally.

Campbell, D. (1969) "Reforms as experiments" *American Psychologist* vol. 24, pp. 409 - 429.

Carroll, J.S. & Johnson, E.J. (1990) *Decision Research: A field guide.* Thousand Oaks, Sage.

Caulley, D.N. (1989) "Towards a short introductory overview of program evaluation" *Bulletin of the Australasian Evaluation Society*, no. 4, May, pp. 14 - 17.

Caulley, D.N. (1993) "Overview of approaches to program evaluation: The Five Generations" in Guthrie, J. (Ed.) *The Australian Public Sector: Pathways to Change in the 1990s.* Sydney, IIR Conferences. pp. 124 - 133.

Caulley, D.N. (1995) "Developing a theory of evaluative criteria" Paper presented at *Evaluation for a New Century: A Global Perspective* (the International Conference of the American Evaluation Association and Canadian Evaluation Society) Vancouver, Canada.

Chalmers, A.F. (1982) *What is this thing called Science? An assessment of the nature and staus of science and its methods*. 2nd ed., University of Queensland Press, St Lucia.

Chamberlin, T. (1965) "The method of multiple working hypotheses" *Science*, vol. 147, pp. 754-759.

Clunie, G., Ludbrook, J. & Faris, I. (1995) "Editorial Comment: The reporting of double-blind, randomized controlled clinical trials and the Cochrane Collaboration" *Australian & New Zealand Journal of Surgery* vol. 65, pp 461 -462.

Cook, T.D. & Campbell, D.T. (1979) *Quasi-experimentation: Design and analysis issues for field settings.* Chicago, Rand-McNally.

Cooper, H.M. (1989) *Integrating Research: A guide for literature reviews.* 2nd ed. Newbury Park, Sage.

Courtis, J.K. (Ed.) 1980 *Research and Methodology in Accounting and Financial Management*. Financial Management Research Centre, Armidale, NSW.

Covey, H.C. (1982) "Basic problems of applying experiments to social programs" *Social Science Review* (September), pp. 424 - 437.

Curnow, W.J. (1995) "Review of Evidence of the Efficacy of Helmets for Cyclists", paper by *Cyclists Rights Action Group* http://www.pcug.org.au/~psvansch/crag/efficacy.htm

Davis, H. R. & Salasin, S.E. (1975) "The Utilisation of Evaluation" In Guttentag, E.L. & Struening, M. (Eds) *Handbook of Evaluation Research*, Vol. 1 Beverly Hills, Sage, pp. 621- 665.

Davis, J.A. (1985) *The Logic of Causal Order.* Sage, Beverly Hills, Ca.

Davis, R. (1992) *Death on the Streets.* Hawes, Nth Yorkshire, Leading Edge Press.

De Cock, G. (1986) "What organizational psychologists say they do and what they really do" in Debus, G. & Schroiff, H. W. (eds.) *The Psychology of Work and Organization: Current Trends and Issues (Proceedings of the West European Conference on the Psychology of Work and Organization* Aachen, F.R.G., 1 - 3 April, 1985*)* Amsterdam, North-Holland, pp. 367 - 378.

De Vries, H., Weijts, W., Dijkstra, M. & Kok, G. (1992) "The utilization of qualitative and quantitative data for health education program planning, implementation, and evaluation: A spiral approach. *Health Education Quarterly* vol. 19, (Spring) pp. 101 - 115.

Dennis, M.L., Fetterman, D.M. & Sechrest, L. (1994) "Integrating qualitative and quantitative evaluation methods in substance abuse research" *Evaluation and Program Planning,* vol. 17, no. 4, pp. 419 - 427.

Department of Finance (1994a) *Doing Evaluation - A Practical Guide.* Canberra, Department of Finance.

Department of Finance (1994b) "Quality Standards in Evaluation " *Evaluation News & Comment*, vol. 3, (no. 2, December) pp. 48.

Department of Finance (1994c) "Quality standards in evaluation" *Evaluation News & Comment*, Vol. 3, no. 2, pp. 48.

Donovan, F. & Jackson, A.C. (1991). *Evaluation and Accountability: Managing Human Service Organisations*. Sydney, Prentice Hall.

Fetterman, D. (1995) "Empowerment Evaluation: An introduction to theory and practice." In Fetterman, D.M., Kaftarian, S.J. & Wandersman, A. (ed.s). *Empowerment Evaluation: Knowledge and tools for Self-Assessment & Accountability*. Thousand Oaks, Ca., Sage.

Funnell, S. (1993) "Reporting the performance of public sector programs". *Evaluation Journal of Australasia,* vol. 5, no. 2, pp. 16 -37

Gardiner, M. & Palmer, G. (1992) *Employment Relations: Industrial Relations and Human Resource Management in Australia* South Melbourne, Mcmillan.

Gross, P.F. (1979) "Recent evaluation activity in Australia". in Senate Standing Committee on Social Welfare, *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 2 Papers Commissioned by the Committee.* Canberra, Australian Government Publishing Service, 1979, p. 58, italics original.

Green, L. W. (1979) "How to evaluate health promotion" *Hospitals*, October, pp. 106 - 108.

Green, L.W. and Ottoson, J.M. 1994. *Community Health*, 7th Edition, McGraw-Hill, St Louis.

Guba E.G. & Lincoln Y.L. (1989) *Fourth Generation Evaluation*. Newbury Park, Sage.

Guba E.G. & Lincoln Y.L. (1990) "Fourth Generation Evaluation: An `interview' with Egon Guba and Yvonna Lincoln" *Evaluation Journal of Australasia*, vol. 2, (3), pp. 3 - 14.

Guttentag, E.L. & Struening, M. (Eds) 1975 *Handbook of Evaluation Research.* vol. 1, Beverly Hills, Sage

Gruenewald, P.J., Treno, A.J., Taff, G. & Klitzner, M. (1997) *Measuring Community Indicators: A systems approach to drug and alcohol problems.* Thousand Oaks, Sage.

Hall, J. (1995) "The practice of economic evaluation in Australia." In Selby-Smith, C. & Drummond, M. (ed.s) *Economic Evaluation in Australian Health Care.* Canberra, AGPS. see http://www.health.usyd.edu.au/people/halljane.htm

Hargreaves, W.A. & Attkisson, C. C. (1978) "Evaluating Program Outcomes". In Attkisson, C. C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds) *Evaluation of Human Service Programs.* N. Y., Academic, pp. 303-339.

Hatry, H.P. (1990) "Performance measurement: A major management tool". In Australasian Evaluation Society (1990) *National Evaluation Conference: Evaluation- You can't manage without it, Volume 1*. Sydney, July. pp. 17-39.

Harrison, M.I. (1994) *Diagnosing Organizations: Methods, Models and Processes.* 2nd ed. Thousand Oaks, Sage.

Hawe, P., Degeling, D., Hall, J. & Brierly, A. (1990) *Evaluating Health Promotion: A Health Worker's Guide.* Sydney, Maclennan + Petty.

Herman, J.L, Morris, L.L. & Fitz-Gibbon, C.T. (1987) *Evaluator's Handbook.* Thousand Oaks, Sage

Herman, J.L., Morris L. & Fitz-Gibbon, C.T. (1987) *The Evaluator's Handbook* vol. 1, *The Program Evaluation Kit*. 2nd.ed., Sage, Beverly Hills, California.

Hollister, R.G. & Hill, J. (1995) "Problems in the evaluation of community wide initiatives" Paper prepared for the *Roundtable on Comprehensive Community Initiatives Russell Sage Foundation Working Paper #70*. New York: Russell Sage Foundation.

Hughes, J. (1980) *The Philosophy of Social Research.* London, Longman.

Hugentobler, M. K., Israel, B.A. & Schurman, S.J. (192) "An action research approach to workplace health: Integrating methods." *Health Eduation Quarterly*, vol. 19 (1) pp. 55 - 76.

Hyman, H. H., Wright, C.R. & Hopkins, T.K. (1962) *Applications of Methods of Evaluation.* Berkeley, Ca. University

of California Press.

Industry Commission (1997) *Report on Government Service Provision: 1997.* Steering Committee for the Review of Commonwealth/State Service Provision, Melbourne, Industry Commission.

Issac, S. & Michael W. B. (1995). *Handbook in Research and Evaluation* 3rd ed.. San Diego, California. Edits.

Joint Committee on Standards for Educational Evaluation (J.R. Sanders, Chair) 1994 *The Program Evaluation Standards: How to Assess Evaluation of Educational Programs.* 2nd ed. Thousand Oaks, Sage.

Kaplan, A. (1964) *The Conduct of Inquiry: Methodology for Behavioral Science.* San Francisco, Chandler.

Kazi, M.A.F. (1996) "Single-case Evaluation in the Public Sector using a combinations of approaches" *Evaluation (The International Journal of Theory, Research and Practice).* vol 2, no. 1, pp. 85 - 97

Kelly, J. R. & McGrath, J.E. (1988). *On time and method*. London, Sage.

Kemmis S. (1985) *The Action Research Planner.* Deakin University, Geelong, Victoria, Australia.

Kimmel, A.J. (1988) *Ethics and Values in Applied Social Research.* Newbury Park, Sage.

Kiresuk, T.J. & Sherman, R.E. 1968 "Goal Attainment Scaling : A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal*, vol. 4, pp. 443-453.

Kiresuk, T.J. & Lund, S.H. (1978) "Goal Attainment Scaling". In Attkisson, C.C., Hargreaves,W.A., Horowitz,M.J. & Sorensen, J.E. (Eds) Evaluation of Human Services Programs. Academic Press, New York, pp. 341-370.

Kiresuk, T.J. & Lund, S.H. (1994) "Implementing Goal Attainment Scaling". In Kiresuk, T.J., Smith, A. & Cardillo, J. E. (Ed.s) *Goal Attainment Scaling: Applications, Theory and Measurement.*.Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Kiresuk, T.J., Smith, A. & Cardillo, J.E. 1994 (Eds.) *Goal Attainment Scaling: Applications, theory, and measurement.* Hillsdale, New Jersey, Lawrence Erlbaum.

Kuhn, T.S. (1970) "Reflections on my critics" In Lakatos, I., & Musgrave, A. (eds) *Criticism and the Growth of Knowledge*. Cambridge University Press, Cambridge, pp. 231-278.

Lakatos, I. (1970) "Falsification and the methodology of scientific research programmes". In Lakatos, I., & Musgrave, A. (eds) *Criticism and the Growth of Knowledge*. Cambridge University Press, Cambridge U.K., pp. 91-196.

Last, J.M.(ed.) 1995. *A Dictionary of Epidemiology*. Oxford University Press, New York.

Leeuw, F.L., Rist, R.C. & Sonnichsen, R.C. (1994) Eds. *Can Governments Learn?: Comparative perspectives on evaluation & organizational learning*. New Brunswick, U.S.A. Transaction Publishers.

Lewis, J.A. Lewis, M.D. & Souflee, F. (1991). *Management of Human Service Programs*. New York, Brooks/Cole Publishing.

Lewis-Beck, M.S. (1993). *Experimental design and methods*. Thousand Oaks, Sage.

Lincoln, Y.S. (1992) "Fourth Generation Evaluation, the Paradigm revolution and health promotion." *Canadian Journal of Public Health -Supplement 1* vol. 83, pp. S6 - S10.

Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*. Beverly Hill, Ca.: Sage.

Linney, J. & Wandersman, A. (1991) *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level - A four-step guide to useful program assessment.* Rockville, Md: Office of Substance Abuse Prevention, US Department of Health & Human Services.

Love, A. (1991) *Internal Evaluation: Building Organizations from Within.* Beverley Hills, Ca: Sage.

Macdonald, G. (1996) "Where next for evaluation ?" *Health Promotion International*, vol. 11, (3), pp. 171 - 173.

MacKay, K. & Robertson, I. (1994) "The influence of evaluation findings in the Commonwealth Budget" *Evaluation News & Comment*, vol. 3, (no. 1, May), pp. 15 - 18.

MacKenzie, R. & Bannister, B. (1988) "Meta-evaluation: Some preliminary thoughts". *The Bulletin of the Australasian Evaluation Society*, No.3, November, pp.19-30.

McKillip, J. (1987) *Need Analysis: Tools for the Human Services and Education.* Newbury Park, Sage.

McLean, I. (1991) "Meta-evaluation - Determining quality and utility". Paper on the On-going Issue: Ethics and Standards in Evaluation, reviewed at the *National Evaluation Conference*, Adelaide, Australasian Evaluation Society and the State Print of S.A., Vol. 2.

Meister, D. (1986) *Human Factors Testing and Evaluation.* Amsterdam, Elsevier. National Agenda for Injury Control, 1992; National Committee for Injury Prevention and Control. 1989

Owen, J.M. (1993). *Program Evaluation. Forms and Approaches*. St Leonards, NSW. Allen & Unwin.

Patton, M.Q. (1987) *How to Use Qualitative Methods in Evaluation.* Newbury Park, California, Sage.

Patton, M.Q. (1990) *Qualitative Evaluation and Research Methods.* Newbury Park, California, Sage.

Patton, M.Q. (1997) *Utilization - focused Evalution: The new century text.* Newbury Park, California, Sage.

Percy-Smith, J. (1996) ed. *Needs Assessment in Public Policy*. Buckingham, UK, Open University.

Platt, J. (1964) "Strong inference" *Science*, vol 146, pp. 347-353.

Rivara, F.P, Beahler, C., Patterson, M.Q., Thompson, D.C. & Zavitkovsky, A. (1997) "Systematic Reviews of Childhood Injury Prevention Interventions: Overview" Harborview Injury Prevention Research Centre, University of Washington, http://weber.u.washington.edu/~hiprc/childinjury/

Robitaille, Y., Legault, J. Abbey, H., & Pless, B. (1990) "Evaluation of an Infant car seat program in a low-income community" *AJDC*, vol. 144, (January) pp. 74 - 78.

Rossi, PH.; & Freeman, HE. (1989). *Evaluation: A Systematic Approach*. 4th Ed. Newbury Park. Sage.

Sackett, D.L., Haynes, R.B. & Tugwell, P. (1985) *Clinical Epidemiology: A basic science for Clinical Medicine*. Boston, Little Brown & Co.

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B. & Richardson, W.S. (1996) "Evidence-based Medicine: What it is and what it isn't" *British Medical Journal*, vol. 312, pp 71 - 72 http://cebm.jr2.ox.ac.uk/ebmisisnt.html

Scriven, M. (1974) "Pros and cons about goal-free evaluation" in Popham, W.J. (ed.) *Evaluation in Education: Current applications.* Berkely, Ca., McCutchan.

Scriven, M. (1991) *Evaluation Thesaurus*, 4th ed., Newbury Park, California, Sage.

Sechrest, L. Stewart, M. Stickle, T.R. & Sidani, S. (1996) *Effective and Persuasive Case Studies. Cambridge, Mass.,* Human Service Research Institute.

Sedgewick S. (1995) "Benchmarking and best practice: promise and performance" *Australian Journal of Public Administration* vol. 54 (September) pp. 401 - 407

Selby, J.V. (1994) "Case-control evaluations of treatment and program efficacy" *Epidemiologic Reviews* vol. 16, no. 1, pp. 90 - 101. Ref # 1994/22

Senate Standing Committee on Social Welfare (1979), *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 1 The Report.* Australian Government Publishing Service, Canberra..

Sharp, C.A. (1994a) "What is Appropriate Evaluation? Ethics and standards in evaluation" *Evaluation News & Comment (*Australasian Evaluation Society), vol. 3, no. 2 (December) pp. 34-41.

Sharp, C.A. (1994b) "Benchmarking: Pros & Cons for evaluation". *WA Branch Conference of the Australasian Evaluation Society*, Perth, 28th September.

Sharp, C.A. (1994c) "Industry best-practice benchmarking in the evaluation context" *Evaluation News & Comment*, vol. 3, no. 1, pp. 27 -33.

Siegel, L. M. Attkisson, C. C. & Carson, L. G. (1978) "Needs identification and program planning in the community context". In Attkisson, C. C. et al. (Eds) *Evaluation of Human Service Programs.* N. Y., Academic, , pp. 213-252.

Silagy, C.A. & Jewell, D. (1994) "Review of 39 years of double-blind, randomized controlled trials in the *British Journal of General Practice*" *British Journal of General Practice*, vol 44, pp. 359- 363.

Slotnick, H.B. (1982) "A simple technique for collecting, analyzing and interpreting evaluative data." *Evaluation and the Health Professions*, vol. 5, no. 3, pp. 245-258.

Smith, M.L. & Glass, G.V. R*esearch and Evaluation in Education and the Social Sciences.* Englewood Cliffs, N.J. Prentice-Hall.

Steering Committee for the Review of Commonwealth/Sate Service Provision (1997) *Report on Government Service Provision.* Volume 1. Melbourne, Australia, Industry Commission/Productivity Commission.

Stake, R.E. (1980) "Program Evaluation, particularly responsive evaluation" in Dockrell, W.B. & Hamilton, D. (eds.) *Rethinking Educational Research* London, Hodder & Stroughton, pp. 72 - 87.

Steckler, A. McLeroy, K.R., Goodman, R.M. & McCormick, L. (1992) "Toward integrating qualitative and quantitative methods: An introduction". *Health Education Quarterly*, vol. 19 (spring) pp. 1 -8.

Stufflebeam, D.L. (1971) "The relevance of the CIPP evaluation model for educational accountability". *Journal of Research and Development in Education*, 5, pp. 19- 25.

Stufflebeam, D.L., Foley, W.J., Gephart, W.J, Guba, E.G., Hammond, R.L. Merriman, H.O. & Provus, M.M. (1971) *Educational Evaluation and Decision Making*. Itasca, Ill.: F.E. Peacock.

Suchman, E.A. (1967) *Evaluative Research: Principles and Practice in Public Service and Social Action Programs.* New York, Russell Sage Foundation.

Susman, G.I. (1984) "Action research: A sociotechnical systems perspective." In G. Morgan (Ed.) *Beyond Method: Strategies for Social Research*. Beverly Hills, Sage, pp. 95 -113.

Van Beurden, E. (1994) "HOOPS (Health Outcome Oriented Problem Segmentation): A graphical framework for health promotion planning, program development and evaluation" *Health Promotion Journal of Australia*, vol. 4 (2) pp. 4 - 8.

Wadsworth, Y. (1984) *Do-it-Yourself Social Research.* Victorian Council of Social Services, & Melbourne Family Care Organisation, Collingwood, Vic., 3066 Australia

Wadsworth, Y. (1991) *Everyday Evaluation on the Run.* Action Research Issues Association (Inc.), 247 -251 Flinders Lane, Melbourne , Vic., 3000, Australia.

Wholey, J.S. (1994) "Evaluability Assessment", *Evaluation: News and Comment*, vol. 3, no. 2, pp. 2-13.

Wholey, J. S., Hatry, H.P. & Newcomer, K.E. (Eds.) 1994 *Handbook of Practical Program Evaluation*. San Fransisco, Jossey-Bass.

Weiss, C.M. (1972) *Evaluation research: Methods for assessing program effectiveness.* Englewood Cliffs, NJ, Prentice-Hall.

Williams, A. P.O. (1986) "Reflections on the interventions of organisational psychologists" in Debus, G. & Schroiff, H. W. (eds.) *The Psychology of Work and Organization: Current Trends and Issues (Proceedings of the West European Conference on the Psychology of Work and Organization* Aachen, F.R.G., 1 - 3 April, 1985*)* Amsterdam, North-Holland, pp. 369 - 385.

Winston, J. (1993) "Performance Indicators: Do they perform?" *Evaluation News & Comment*, vol. 2, (2): 22 - 39

Yin R.K. (1994) *Case study research: design and methods*. 2nd ed. Thousand Oaks, CA: Sage Publications.

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

REFERENCES ON EVALUATION

SECTION 1: OVERVIEW REFERENCES

There are three types of reference categorized in the overview bibliography, viz:

A. Useful texts
B. Introductory Handbooks
C. Specialised Technical Guides
D. Internet or WWW resources

A. Useful Texts

Attkisson, C.C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds). 1978. *Evaluation of Human Service Programs.* N. Y., Academic.

Donovan, F. & Jackson, A.C. (1991). *Evaluation and Accountability: Managing Human Service Organisations*. Sydney, Prentice Hall.

Gay, L.R. (1980) *Educational Evaluation & Measurement: Competencies for analysis and application.* Columbus, Oh., Charles E. Merrill.

Guttentag, E.L. & Struening, M. (Eds) 1975 *Handbook of Evaluation Research.* vol. 1, Beverly Hills, Sage

Hawe, P., Degeling, D., Hall, J. & Brierly, A. (1990) *Evaluating Health Promotion: A Health Worker's Guide.* Sydney, Maclennan + Petty.

Lewis, J.A. Lewis, M.D. & Souflee, F. (1 991). *Management of Human Service Programs*. New York, Brooks/Cole Publishing.

Love, A. (1991) *Internal Evaluation: Building Organizations from Within.* Beverley Hills, Ca: Sage.

Owen, J.M. (1993). *Program Evaluation. Forms and Approaches*. St Leonards, NSW. Allen & Unwin.

Rossi, PH.; & Freeman, HE. (1989). *Evaluation: A Systematic Approach*. 4th Ed. Newbury Park. Sage.

Scriven, M. (1991) *Evaluation Thesaurus*, 4th ed., Newbury Park, California, Sage.

B. Introductory Handbooks

Australian Youth Foundation & Sharp, C.A. (1996) *Do-it-yourself Evaluation Manual: An integrated approach to project management and evaluation.* Sydney. Australian Youth Foundation.

Donovan, F. & Jackson, A.C. (1991). *Evaluation and Accountability: Managing Human Service Organisations*. Sydney, Prentice Hall.

Hawe, P., Degeling, D., Hall, J. & Brierly, A. (1990) *Evaluating Health Promotion: A Health Worker's Guide.* Sydney, Maclennan + Petty.

Herman, J.L, Morris, L.L. & Fitz-Gibbon, C.T. (1987) *Evaluator's Handbook.* Thousand Oaks, Sage.

Isaac, S. & Michael, W. B. (1995) *Handbook in Research and Evaluation.* 3rd ed. San Diego, Ca: Edits.
Linney, J. & Wandersman, A. (1991) *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level - A four-step guide to useful program assessment.* Rockville, Md: Office of Substance Abuse Prevention, US Department of Health & Human Services.
Kemmis S. (1985) *The Action Research Planner.* Deakin University, Geelong, Victoria, Australia.
Lewis, J.A. Lewis, M.D. & Souflee, F. (1991). *Management of Human Service Programs*. New York, Brooks/Cole Publishing.
Wadsworth, Y. (1984) *Do-it-Yourself Social Research.* Victorian Council of Social Services, & Melbourne Family Care Organisation, Collingwood, Vic., 3066.
Wadsworth, Y. (1991) *Everyday Evaluation on the Run.* Action Research Issues Association (Inc.), 247 -251 Flinders Lane, Melbourne , Vic., 3000.
Wholey, J. S., Hatry, H.P. & Newcomer, K.E. (Eds.) 1994 *Handbook of Practical Program Evaluation*. San Fransisco, Jossey-Bass.

C. Specialised Technical Guides

Agar, M.H. (1986) *Speaking of ethnography*. London, Sage Publications.
Allison, P.D. (1984) *Event history analysis: regression for longitudinal event data*. London, Sage.
Bogden, R. & Knopp B.S. (1992). "Foundations of qualitative research in education." In Qualitative Research for Education: An Introduction to Theory and Methods (2nd ed.). Boston. Allyn and Bacon.<
Boruch, R.F. (1997) *Randomized experiments for planning and evaluation: a practical guide*, London, Sage.
Brinkerhoff, R.O. & Dressler, D.E. (1990). *Productivity measurement: a guide for managers and evaluators*. London, Sage Publications.
Dillman, D.A. (1991). "The Design and Administration of Mail Surveys."*Annual Review Sociology*, vol. 17, pp. 225 - 249
Edwards, W. & Newman, J.R. (1982) *Multiattribute Evaluation*. Newbury Park, Sage.
Fielding, N.G; and Fielding, J., (1986). *Linking data*. London, Sage.
Feldman, M.S. (1995) *Strategies for interpreting qualitative data*. London, Sage.
Fowler, F. J. Jnr. & Mangione, T.W. 1990. *Standardized survey Interviewing: minimizing interviewer-related error*. London, Sage
Fowler, F. J. Jnr. (1995). *Improving survey questions: design and evaluation*. London, Sage Publications.
Gubriam, J. 1988. *Analyzing field reality*. London, Sage.
Hamel, J., Dufour, S. & Fortin, D. (1993) *Case study methods*. London, Sage.
Harrison, M.I. 1994. *Diagnosing organizations: methods, models, and processes*. London, Sage Publications.
Kimmel, A.J. 1988. *Ethics and values in applied social research*. London, Sage
Lavrakas, P.J. (1993). *Telephone survey methods: sampling, selection, and supervision*. London, Sage Publications.
Lewis-Beck, M.S. (1993). *Experimental design and methods*. London, Sage
McDowall, D., McCleary, R., Meidinger, E.E. & Hay, R.A. Jr.(1980) *Interrupted time series analysis*. London, Sage.
McKillip, J. 1987. *Need analysis: tools for the human services and education*. London, Sage Publications
Morgan, D. L. (1997) *Focus groups as qualitative research*. London, Sage.
Noblit, George W. and Hare, R Dwight.1988. *Meta-ethnography: synthesizing qualitative studies*. London, Sage Publications.
Weller, S.C.; and Romney, A K., (1988) *Systematic data collection.* London, Sage.
Wolcott, H.F. 1990. *Writing up qualitative research*. London, Sage.

D. Philosophical and Methodological Issues in Evaluation

Boruch, R.F. & Boe, E. (1995) 'The use of science and mathematics education indicators and studies: A briefing' in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 7- 13, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Chelimsky, E. (1995) Commentary. in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 127 - 128, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Chubin, D.E. (1995) 'Foreword' in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* p. v, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Fetterman, D. (1995) "Empowerment Evaluation: An introduction to theory and practice." In Fetterman, D.M. Kaftarian, S.J. & Wandersman, A. *Empowerment Evaluation: Knowledge and tools for Self-Assessment &*

Accountability. Thousand Oaks, Ca., Sage.
Frechtling, J.A., Ed., (1995) *Footprints: Strategies for Non-Traditional Program Evaluation.* Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA., pp. 7- 13.
Johnson, S. T. (1995) 'Searching near, far, and wide: A plan for evaluation' in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 15 - 23, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Kiresuk, T.J. (1994) "Historical perspective" in Kiresuk, T.J. Smith, A. & Cardillo, J.E. (ed.s) *Goal Attainment Scaling: Applications, Theory and Measurement* pp. 135 - 160, Hillsdale New Jersey: Lawrence Erlbaum,
Sechrest, L. Stewart, M. Stickle, T.R. & Sidani, S. (1996) *Effective and Persuasive Case Studies. Cambridge, Mass.,* Human Service Research Institute.
Schick, A. (1996) "The spirit of reform: Managing the NZ State Sector in a time of change" Wellington, New Zealand, State Services Commission. (http://www.ssc.govt.nz/documents/reform1.htm)
Schwandt, T. & Halpern, E. S. (1988) *Linking Auditing and Metaevaluation: Enhancing Quality in Applied Research.* (Applied Social Research Methods Series Volume 11), Newbury Park, Cal., Sage.
Scriven, M. (1972) "Prose and cons about gaol-free evaluation" *Evaluation Comment: The Journal of Educational Evaluation* vol 3. (4), pp. 1- 7.
Scriven, M. (1991) *Evaluation Thesaurus.* 4th ed., Newbury Park, Ca., Sage.
Scriven, M. (1995) "Overview" in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 131 - 138, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Sharp, C.A. (1995) "Evaluation as a tool of empowerment." *Consuming Interest.* (Australian Consumers' Association), No. 62, Summer, 1995, p 27.
Sharp, C.A. (1991) 'The Program evaluation Kit, 2nd edition.' Book review *Evaluation Journal of Australasia*, 3 (1): pp 46 - 50.
Sharp, C.A., Roffey, B.L. & Lewis, N.R. (1993) 'Appropriateness in Audit and Evaluation: "It all depends, Minister!"' *Accounting, Finance and Management Research Paper 3/93*, School of Coomerce, Flinders University of South Australia.
Sharp, C.A. (1994) 'What Is Appropriate Evaluation? Ethics and standards in evaluation'. *Evaluation News & Comment (*Australasian Evaluation Society), 3 (2): 34-41.
Sharp, C.A. (1991) "The Program evaluation Kit, 2nd edition." Book review *Evaluation Journal of Australasia*, vol. 3, No. 1, pp 46 - 50.
Sharp, C.A. (1994) "What Is Appropriate Evaluation? Ethics and standards in evaluation" *Evaluation News & Comment (*Australasian Evaluation Society), vol. 3, no. 2 (December) pp. 34 - 41.
Sharp, C.A. (1994) "Industry best-practice benchmarking in the evaluation context" *Evaluation News & Comment (*Australasian Evaluation Society), vol. 3, no. 1, pp. 27 -33.
Sharp, C.A. (1996a) *Do-it-Yourself Evaluation Manual.* Sydney, Australian Youth Foundation.
Sharp, C.A. (1996b) "Footprints: Strategies for Non-Traditional Program Evaluation" *Evaluation: The International Journal of Theory, Research and Practice*, vol. 2, no. 2, pp. 226 - 231.
Stake, R.E. (1995) "The virtual reality of systematic effects of NSF programming on education: Its profession, practice, research and institutions" in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 107 - 125, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Webb, N. L. (1995) "Communicating the value of the National Science Foundation's contributions to research and innovative technical applications for mathematics and science education." in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 53 - 74, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.
Webb, E. T., Campbell, D.T., Schwartz, Sechrest, L. & Grove, J.B. (1981) *Nonreactive Measures in the Social Sciences.* 2nd ed.Boston: Houghton Mifflin.
Yin, R.K. (1995) 'New methods for evaluating programs in NSF's Division of Research, Evaluation and Dissemination' in Joy A. Frechtling (ed.) *Footprints: Strategies for Non-Traditional Program Evaluation.* pp. 25 - 36, Directorate for Education and Human Resources, National Science Foundation, Arlington Va., USA.

E. SPECIFIC EVALUATION TECHNIQUES

Needs Analysis & Planning

Fagence, M. *Citizen Participation in Planning*. Oxford: Pergamon, 1977.
Green, L.W., Kreuter, M.W., Deeds, S.G., & Partridge, K.B. (1980) *Health Planning - A Diagnostic Approach.* Palo Alto, Mayfield Publishing.
Green, L.W. & Anderson, C.L. (1986).*Community Health.* St Louis, Times Mirror/Mosby College Publishing.

Green, L.W. & Kreuter, M.W. (1991) *Health Promotion Planning: An educational and Environmental Approach.* Pale Alto, Mayfield Publishing.
Isaac, S. & Michael, W.B. (1995) *Handbook in Research and Evaluation*. 3rd ed. San Diego: Edits.
Percy-Smith, J. (1996) ed. *Needs Assessment in Public Policy*. Buckingham, UK, Open University.
Scriven, M. (1972) "Prose and cons about goal-free evaluation" *Evaluation Comment: The Journal of Educational Evaluation* vol 3. (4), pp. 1- 7.
Scriven, M. (1974) "Prose and cons about goal-free evaluation" In W.J. Popham (ed.) *Evaluation in Education: Current applications.* Berkely, Ca. McCutchan.
Scriven, M. (1991) *Evalution Thesaurus* 4th edition, Newbury Park, Sage.
Siegel, L. M. Attkisson, C. C. & Carson, L. G. Needs identification and program planning in the community context. In Attkisson, C. C. et al. (Eds) *Evaluation of Human Service Programs.* Academic, N. Y., 1978, pp. 213-252.

Group Processes

Delbecq, A. & Van de Ven, A. H. (1971) "A group process model for problem identification and program planning" *Journal of Applied Behavioral Science* Vol. 7, No. 4, pp. 466 - 492.
Fagence, M. (1977) *Citizen Participation in Planning.* Oxford: Pergamon.

Ethics and Standards in Evaluation

Joint Committee on Standards for Educational Evaluation, (1994) *The Program Evaluation Standards*. 2nd edition, Thousand Oaks, California, Sage.
Sharp, C.A. (1994) "What is appropriate evaluation? Ethics and Standards in Evaluation" *Evaluation News & Comment*, Vol. 3, no. 2, pp. 34 - 41. (Published by the Australasian Evaluation Society, see below)
Caulley, D. (1994a) "The program evaluation standards (2nd edition)" *Evaluation News & Comment*, Vol. 3, no. 2, pp. 42 - 44.
Caulley, D. (1994b) "Guiding Principles for Evaluators from the American Evaluation Association" *Evaluation News & Comment*, Vol. 3, no. 2, pp. 45 - 47.

Evaluability Assessment

Davis, H. R. & Salasin, S.E. (1975) "The Utilisation of Evaluation" In Guttentag, E.L. & Struening, M. (Eds) *Handbook of Evaluation Research.*, Vol. 1 Beverly Hills, Sage, pp. 621-665.
Kiresuk, T.J. & Lund, S.H. (1994) "Implementing Goal Attainment Scaling". In Kiresuk, T.J., Smith, A. & Cardillo, J. E. (Ed.s) *Goal Attainment Scaling: Applications, Theory and Measurement.*.Hillsdale, New Jersey, Lawrence Erlbaum Associates.
Wholey, J.S. (1994) "Evaluability Assessment" *Evaluation: News & Comment*, vol. 3, no. 2 pp. 2-13

Project logic designs

Funnell, S. (1997). 'Program Logic: An adaptable tool for designing and Evaluating programs' *Evaluation: News and Comment*, vol 6, no 1, July, pp 5-17
Lenne, B. & Cleland, H. (1987) "Describing Program Logic". *Program Evaluation Bulletin*, No. 2, Program Evaluation Unit Public Service Board of New South Wales, Sydney.
Linney, J.A. & Wandersman, A. (1991). *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level - A Four-Step Guide to Useful Program Assessment.*Office of Substance Abuse Prevention, U.S. Department of Health & Human Services, 5600 Fishers Lane, Rockwall II, Rockville, Maryland, USA, 20857
Rowe, A. (1995) *The Four-Step Assessment: Monitoring material for projects supported by the Canada/Newfoundland Cooperation Agreement on Human Resources Development.* Andy Rowe Consultants, P.O. Box, 155 Stn C, St John's Newfoundlan, Canada, A1C 5J2.
Rustomji, L. (1987) "Reporting public sector performance". *Program Evaluation Bulletin*, no. 3. Sydney, Program Evaluation Unit, Public Service Board of New South Wales.

Other References on Evaluation Design:

Fitz-Gibbon, C.T. & Morris, L.L. (1978) *How to Design a Program Evaluation* (*Program Evaluation Kit,* Volume 3) Beverly Hills, California, Sage.
Stecher, B. & Davis, W.A. (1987) *How to Focus an Evaluation* (Program Evaluation Kit Volume 2) Beverly Hills, California, Sage.

Objectives Oriented (featuring Goal Attainment Scaling)

Bloom, B.S., Hastings, J.T. & Madaus, G.F. (1971) *Handbook on Formative and Summative Evaluation of Student Learning.*New York, McGraw-Hill.
Kiresuk, T.J. & Lund, S.H. (1978) "Goal Attainment Scaling". In Attkisson, C.C., Hargreaves,W.A., Horowitz,M.J. &

Sorensen, J.E. (Eds) *Evaluation of Human Services Programs*. Academic Press, New York, pp. 341-370.
Kiresuk, T.J. & Lund, S.H. (1994) "Implementing Goal Attainment Scaling". In Kiresuk, T.J., Smith, A. & Cardillo, J. E. (Ed.s) *Goal Attainment Scaling: Applications, Theory and Measurement.*.Hillsdale, New Jersey, Lawrence Erlbaum Associates.
Mager, R.F. (1962) *Preparing Instructional Objectives.* Palo Alto, Ca., Fearon.

Other References:
Hargreaves, W.A. & Attkisson, C. C. (1978) "Evaluating Program Outcomes". In Attkisson, C. C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds) *Evaluation of Human Service Programs.* Academic, N. Y., 1978, pp. 303-339.

Satisfaction evaluation

Slotnick, H.B. (1982) "A simple technique for collecting, analyzing and interpreting evaluative data". *Evaluation and the Health Professions*, vol. 5, no. 3, pp. 245-258.

Other References
Harrison, S.R. & Tamaschke, H.U. (1987) *Appled Statistical Analysis*. Syndey, Prentice-Hall.
Henerson, M.E., Morris,L.L. & Fitz-Gibbon, C.T. (1987) *How to Measure Attitudes (The Program Evaluation Kit,* Volume 6), 2nd Edition, Beverly Hills, California, Sage.
King, J.A. Morris, L. L. & Fitz-Gibbon, C.T. (1987) *How to Assess Program Implementation (The Program Evaluation Kit,* Volume 5), 2nd Edition, Beverly Hills, California, Sage.
Morris,L.L.; Fitz-Gibbon, C.T. & Lindheim, E.(1978) *How to Measure Performance and Use Tests (The Program Evaluation Kit,* Volume 7), 2nd Edition, Beverly Hills, California, Sage.
Morris,L.L. & Fitz-Gibbon, C.T. (1978) *How to Analyze Data (The Program Evaluation Kit,* Volume 8), 2nd Edition, Beverly Hills, California, Sage.
Moser, C.A. & Kalton, G. (1981) *Survey Methods in Social Investigation*. 2nd edition., London, Heinemann Educational Books.
Patton, M.Q. (1987) *How to Use Quantitative Methods in Evaluation (The Program Evaluation Kit,* Volume 4), 2nd Edition, Beverly Hills, California, Sage.

Reporting

Department of Finance (1994) "Quality standards in evaluation" *Evaluation News & Comment*, Vol. 3, no. 2, pp. 48.
Funnell, S. (1993) "Reporting the performance of public sector programs". *Evaluation Journal of Australasia,* vol. 5, no. 2, pp. 16 -37
Morris, L.L., Fitz-Gibbon, C.T. & Freeman, M.E. (1987) *How to Communicate Evaluation Findings* (*The Program Evaluation Kit,* Volume 9) , 2nd Edition, Beverly Hills, California, Sage.

F. Basic Guides to Evaluation:

Abbott, D. J. & Craig, B. (1994) *How to Evaluate Your Community Health Programs.* Revised edition, Adelaide, SA, South Australia Community Health Research Unit.
CWAV Packages in Quality and Accountability. ***A Guide to Standards for Community Services****.* Melbourne: Children's Welfare Association of Victoria. (1986-1987), including:
-A Guide to Establishing Accountability.
-A Guide to Basic Goal Setting, Monitoring and Evaluation.
-A Guide to Measuring Performance and Outcomes.
-A Guide to Agency Review.
-A Guide to Negotiation, Advocacy and Public Relations.
-A Guide to Employing a Consultant
-A Guide to Basic Information Systems.
Linney, J.A. & Wandersman, A. *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level - A Four-Step Guide to Useful Program Assessment.*Office of Substance Abuse Prevention, U.S. Department of Health & Human Services, 5600 Fishers Lane, Rockwall II Rockville, Maryland, USA, 20857
Sharp, C.A. (1996) *Do-it-yourself Evaluation Manual: An integrated approach to project management and evaluation.* Sydney. Australian Youth Foundation.
Wadsworth, Y. (1984) *Do-it-Yourself Social Research.* Victorian Council of Social Services, & Melbourne Family Care Organisation, Collingwood, Vic., 3066.
Wadsworth, Y. (1991) *Everyday Evaluation on the Run.* Action Research Issues Association (Inc.), 247 -251

Flinders Lane, Melbourne , Vic., 3000.
Wholey, J. S., Hatry, H.P. & Newcomer, K.E. (Eds.) 1994 *Handbook of Practical Program Evaluation*. San Fransisco, Jossey-Bass.

G. EVALUATION OF INJURY PREVENTION

Dannenberg, A.L. (1996) "Evaluation of Injury Interventions: An Overview." Symposium Session 12 in *Abstracts of the Third International Conference on Injury Prevention and Control,* 18 - 22 February, Melbourne, NISU, Australian Institute of Health & Welfare, Commonwealth Department of Human Services and Health, p. 1.
Dannenberg, A.L. & Fowler, C.J. (1997) "Evaluation of Interventions to Prevent Injuries: An Overview" Unpublished Personal Communication (August, 1997), Center for Injury Research and Policy, Johns Hopkins University School of Hygiene and Public Health.
Harrison, J.E. & Cripps, R.A. (1994) ed.s *Injury in Australia: An epidemiological review.* Department of Human Services and Health, Canberra, AGPS.
Miller, T.R. (1996) "Tools for Evaluating Injury Prevention: Benefits and cost"Analysis. Symposium Session 12 in *Abstracts of the Third International Conference on Injury Prevention and Control,* 18 - 22 February, Melbourne, NISU, Australian Institute of Health & Welfare, Commonwealth Department of Human Services and Health, p. 1.
Rivara, F.P, Beahler, C., Patterson, M.Q., Thompson, D.C. & Zavitkovsky, A. (1997) "Systematic Reviews of Childhood Injury Prevention Interventions: Overview" Harborview Injury Prevention Research Centre, University of Washington,
Towner, E.M.L. (1996) "Evaluation of Educational Interventions"Symposium Session 12 in *Abstracts of the Third International Conference on Injury Prevention and Control,* 18 - 22 February, Melbourne, NISU, Australian Institute of Health & Welfare, Commonwealth Department of Human Services and Health, p. 1.
Zwerling, C., Daltroy, L.H., Fine, L.J., Johnston, J.J., Melius, J. &. Silverstein, B.A. (1997) "Design and Conduct of Occupational Injury Intervention Studies: A Review of Evaluation Strategies" *American Journal of Industrial Medicine vol.* 32, pp.164-179

H. Internet or WWW Resources

These World Wide Web and internet/email addresses may be repeated in any of the following sections, as well as in the text and other associated reference lists.

Specific Injury Prevention & Health Promotion WWW Sites & Email addresses
AUSTRALIAN

NISU Web Page for the FIPPM Evaluation Annotated Bibliography Project
http://www.nisu.flinders.edu.au/nnnisu/evalcall.html
see also: NISU (Australian National Injury Surveillance Unit)
http://www.nisu.flinders.edu.au/
Including:
National Data Standards for Injury Surveillance (NDSIS),
http://www.nisu.flinders.edu.au/pubs/manuals/ndsis/ndsisman.html
Australian Injury Prevention Bulletin Issue number 15: "Progress and current issues in child injury prevention" by Jerry Moller and Renate Kreisfeld
http://www.nisu.flinders.edu.au/pubs/bulletin15/bulletin15-Data.html
Centre for Health Economics Research and Evaluation (CHERE)
University of Sydney, Camperdown, NSW, 2050 Australia
http://www.chere.usyd.edu.au/
email:chere@pub.health.usyd.edu.au
Centre for Health Program Evaluation (CHPE)
Deparment of Public Health anbd Community Medicine,
University of Melbourne &
Health Economics Unit, Faculty of Business, Monash University
Austin & Repatriation Medical Centre
West Heidelberg, Vic., 3081
http://ariel.ucs.unimelb.edu.au/~chpe/
email:CHPE@BusEco.monash.edu.au

INTERNATIONAL

US National Centre for Injury Prevention and Control
http://www.cdc.gov/ncipc/ncipchm.htm
Has 10 Research Centres and see the following Associated Internet Addresses re Injury See whos-who site:
http://www.edc.org/HHD/csn/buildbridges/whoswho/ncipc.html and
http://www.injurycontrol.com/icrin/
http://www.sph.emory.edu/CIC/cichome.html
http://weber.u.washington.edu/~hiprc/
http://www.cdc.gov/ncipc

National Highway Traffic Safety Administration (NHTSA).
Building Bridges seeks to promote collaboration between public health and traffic safety professionals to reduce the toll that motor vehicle-related injuries take upon our society. Building Bridges, a project of *Education Development Center*, Inc., is funded by the National Highway Traffic Safety Administration (NHTSA).
http://www.edc.org/HHD/csn/buildbridges/index.html

Systematic Reviews of Childhood Injury Prevention Interventions
http://weber.u.washington.edu/~hiprc/childinjury

The Injury Control Resource Information Network
http://www.injurycontrol.com/icrin

P. Phillipe's Web Site, PHD; Social and Preventive Medicine, University of Montreal, Quebec, Canada.
http://alize.ere.umontreal.ca/~philippp

Specific Evaluation WWW Sites & Email addresses

Australasian Evaluation Society (AES)
PO Box 448
Curtin, A.C.T., 2605
Australia
Telephone: 06- 282 3320
Fax: 06- 282-3058
Email: aesoffic@ozemail.com.au
WWW Homepage: http://www.parklane.com.au/aes

A.E.A (American Evaluation Association)
401 East Jefferson St., Suite 205
Rockville MD USA 20850
*Email: **aea@phoenixpp.com***
Ph: 301 251 7700

General: Government Agencies in Australia
http://www.nla.gov.au/oz/gov/ or
http://www.dfat.gov.au/
This lists several Australian Government agencies and their WWW home pages for you to follow up. If you want to know about evaluation in Government then try the Department of Finance's list of evaluation reports and online publications:

Australian Commonwealth Dept of Finance:
http://www.dofa.gov.au/
The Department of Finance has responsibility for (financial) performance management improvement (including benchmarking and contestability) in the Australian Government, through the work of the Management Advisory Board (MAB) and the Management Improvement Advisory council (MIAC). Steering Committee for the Review of Commonwealth/Sate Service Provision (1997) *Report on Government Service Provision.*

American Evaluation Association
http://www.eval.org/

Canadian Evaluation Society
http://www.unites.uqam.ca/sce/ces-sce.html

Other Related WWW Sites & Email addresses

Australian Institute of Health & Welfare
http://www.aihw.gov.au/

Australian Department of Health & Aged Care
http://www.health.gov.au/pubhlth/strateg/injury/index.htm
InterNet Health Resources
http://www.arcade.uiowa.edu/hardin-www/md.html (for text-based users)
MedWeb:Public Health
http://www.cc.emory.edu/WHSCL/medweb.ph.html
The Global Health Network
http://globalhealth.pitt.edu
The World-Wide Web Virtual Library: Epidemiology
http://www.epibiostat.ucsf.edu/epidem/epidem.html
Canadian Society for International Health
http://www.csih.org/csihmem.html
Qualitative Software Discussion Group:
This discussion group is the creation of the CAQDAS (ESRC funded) Project. This was set up to disseminate information and understanding about software developed to assist analysis and handling of qualitative data, (That is any kind of textual data, eg documents, interview transcripts, field notes etc.) in the behavioural and social sciences. More information can be found at:
http://www.mailbase.ac.uk/lists/qual-software/welcome.html
The National Centre for Health Statistics (USA)
http://www.cdc.gov/nchswww/nchshome.htm
Epidemio-l Archives (a service offered by Matthias Wjst)
http://pc2109.gsf.de/epidemio-l/search.cfm
Sage Publications Monographs
http://www.sagepub.com
Victorian Department of Human Services
http://hna.ffh.vic.gov.au/yafs/index.html
http://home.vicnet.net.au/~saapmelb/
http://home.vicnet.net.au/~saapmelb/family.htm#_family
Internet Addresses re (Bio)Statistics
http://www.biostat.washington.edu/Xvlib/
http://www.ams.med.uni-goettingen.de/~rhilger/ListS_B.html
http://www-personal.umich.edu/~dronis/statfaq.htm
http://www.mailbase.ac.uk/lists-k-o/minitab/files/
http://www.minitab.com/
http://www.stattransfer.com/lists.html
http://www.stats.gla.ac.uk/allstat/index.html
http://usd-inc.com/pepi.html
http://www.epibiostat.ucsf.edu/epidem.html
http://www.healthworks.co.uk/hw/publisher/arnold/arnold11.html
http://www.maths.uq.oz.au/~gks/webguide/journals.html

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA
Research Centre for Injury Studies
Incorporating the AIHW National Injury Surveillance Unit

References Cited in the Annotated Bibliography

NOTE: In keeping with the general format, the references which have been annotated are listed here in alphabetical order.

Ref # 1978/2 Attkisson, C. C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds) Evaluation of Human Service Programs. N.Y., Academic.

Ref # 1987/1 Baum, F. & Brown, V. (1987) Healthy Cities Australia: From vision to reality. Sydney: Australian Community Health Association.

Ref # 1994/24 Berry, TD; Gilmore, MR; & Geller, ES. 1994. "An Individual-Subject Approach to the Study of Community-Based Interventions". Environment and Behavior. Vol 26(4), pp 451-476.

Ref # 1991/1 Bowen, GL., Farkas, G. & Neenan, PA. 1991. "Application of Time-Series Designs to the Evaluation of Social service Program Initiatives: The Recycling Fund Concept". Social Work Research & Abstracts. Vol 27(3), pp 9-15.

Ref # 1993/1 Bradley, C. 1993. "Designing Medical and Educational Intervention Studies: A Review of Some Alternatives to Conventional Randomized Control Trials". Diabetes Care. Vol 16(2), pp 509-518.

Ref # 1994/2 Clarke, JA. & Langley, JD. 1994. "Firework Related Injury in New Zealand". The New Zealand Medical Journal. Vol 107, pp 423-425.

Ref # 1993/2 Corcoran, KJ. 1993. "Practice Evaluation: Problems and Promises of Single-System Designs in Clinical Practice". Journal of Social Service Research. Vol 18, No. 1(2), pp 147-159.

Ref # 1997/5 Day, L., Ozanne-Smith, J., Cassell, E. & McGrath, A. 1997. Latrobe Valley Better Health Project: Evaluation of the Injury Prevention Program 1992-1996. Monash University Accident Research Centre. Report # 114. http://www.general.monash.edu.au/MUARC/rptsum/es114.htm

Ref # 1993/3 Dehar, M., Casswell, S. & Duignan, P. 1993. "Formative and Process Evaluation of Health Promotion and Disease Prevention Programs". Evaluation Review. Vol 17(2), pp 204-220.

Ref # 1994/1 Dennis, ML., Fetterman, DM. & Sechrest, L. 1994. "Integrating Qualitative and Quantitative Evaluation Methods in Substance Abuse Research". Evaluation and Program Planning. Vol 17(4), pp 419-427.

Ref # 1992/1 Desai, A. & You, M. 1992. "Policy Implications From an Evaluation of Seat Belt Use Regulation". Evaluation Review. Vol 16(3), pp 247-265.

Ref # 1997/4 Elvik, R. 1997. "Evaluations of Road Accident Blackspot Treatment: A Case of the Iron Law of Evaluation Studies". Accident Analysis and Prevention. Vol 29 (2), pp 191-199.

Ref # 1995/4 Fetterman, D. (1995) "Empowerment Evaluation: An introduction to theory and practice." In Fetterman, D.M. Kaftarian, S.J. & Wandersman, A. Empowerment Evaluation: Knowledge and tools for Self-Assessment & Accountability. Thousand Oaks, Ca., Sage.

Ref # 1988/3 Fisher, L. 1988. "Childhood Injuries - Causes, Preventive Theories and Case Studies". Journal of Environmental Health. Vol 50, pp 355-360.

Ref # 1992/3 Gielen, AC. (1992). "Health Education and Injury Control: Integrating Approaches". Health Education Quarterly. Vol 19(2), pp 203-218.

Ref # 1994/10 Goldenhar, LM. & Schulte, PA. 1994. "Intervention Research in Occupational Health and Safety". Journal of Occupational Medicine. Vol 36(7), 763-775.

Ref # 1996/4 Goldenhar, LM. & Schulte, PA. 1996. "Methodological Issues for Intervention Research in Occupational Health and Safety". American Journal of Industrial Medicine. Vol 29, pp289-294.

Ref # 1979/1 Green, L.W. (1979) "How to evaluate health promotion" Hospitals, October, pp. 106 - 108.

Ref # 1980/1 Green, L.W. Kreuter, MW., Deeds, S.G., & Partridge, K.B. (1980) Health Eduation Planning -A Diagnostic Approach. Palo Alto, Mayfield Publishing.

Ref # 1989/3 Guba, E.G. & Lincoln, Y. S. (1989); Fourth Generation Evaluation. Newbury Park, Sage.

Ref # 1990/4 Hawe, P., Degeling, D., Hall, J. & Brierly, A. (1990) Evaluating Health Promotion: A Health Worker's Guide. Sydney, Maclennan + Petty.

Ref # 1987/2 Herman, J.L, Morris, L.L. & Fitz-Gibbon, C.T. (1987) Evaluator's Handbook. Thousand Oaks, Sage.

Ref # 1995/1 Hollister, R.G. & Hill, J. (1995) "Problems in the evaluation of community wide initiatives" Paper prepared for the Roundtable on Comprehensive Community Initiatives Russell Sage Foundation Working Paper No. 70. New York: Russell Sage Foundation; http://epn.org/sage/rsholl.html

Ref # 1992/4 Hugentobler, M. K., Israel, B.A. & Schurman, S.J. (1992) "An action research approach to workplace health: Integrating methods." Health Education Quarterly, vol. 19 (1) pp. 55 - 76.

Ref # 1978/1 Kiresuk, T.J. & Lund, S.H. (1978) "Goal Attainment Scaling". In Attkisson, C. C.; Hargreaves, W. A.; Horowitz, M. J. and Sorensen, J. E.. (Eds) Evaluation of Human Service Programs. Academic, N. Y., 1978.

Ref # 1988/1 Langley, J. & McLoughlin, E. 1988. "Difficulties and Bonuses of Evaluation: Evaluating New Zealand's Children's Nightclothes Act 1977". Burns Including Thermal Injuries. Vol 14(6), pp 435-439.

Ref # 1990/3 Layde, PM. 1990. "Beyond Surveillance: Methodologic Considerations in Analytic Studies of Agricultural Injuries". American Journal of Industrial Medicine. Vol 18(2), pp 193-200.

Ref # 1985/1 Lincoln, Y.S. & Guba, E.G. (1985) Naturalistic Inquiry. Beverly Hill, Ca.: Sage.

Ref # 1996/5 Macdonald, G. (1996) "Where next for evaluation ?" Health Promotion International, vol. 11, (3), pp. 171 - 173.

Ref # 1995/2 Marconi, KM., & Rudzinski, KA. 1995. "A Formative model to Evaluate Health Sevices Research". Evaluation Review. Vol 19(5), pp 501-510.

Ref # 1986/2 Meister, D. (1986) Human Factors Testing and Evaluation. Amsterdam, Elsevier.

Ref # 1995/3 Oakley, A., Fullerton, D., Holland, J., Arnold, S., France-Dawson, M., Kelley, P., & McGrellis, S. 1995. "Sexual Health Education Interventions for Young People: A Methodological Review". BMJ. Vol 310, pp 158-162.

Ref # 1994/18 Orwin, RJ., Sonnefeld, LJ., Garrison-Mogren, R, & Gray Smith, N. 1994. "Pitfalls in Evaluating the Effectiveness of Case Management Programs for Homeless Persons: Lessons From the NIAAA Community Demonstration Program". Evaluation Review. Vol 18(2), pp 153-207.

Ref # 1995/3 Province, MA., Hadley, EC., Hornbrook, MC., Lipsitz, LA., Miller, JP., Mulrow, CD., Ory, MG., Sattin, RW., Tinetti, ME. & Wolf, SL. 1995. "The Effects of Exercise on Falls in Elderly Patients: A Preplanned Meta-Analysis of the FICSIT Trials". JAMA. Vol 273, pp 1341-1347.

Ref # 1996/1 Raycovich, KST., McCormick, MC., Howell, EM. & Devaney BL. 1996. "Evaluating the Health Start Program: Design Development to Evaluative Assessment". Evaluation & the Health Professions. Vol 19(3), pp 342-362.

Ref # 1990/2 Robitaille, Y., Legault, RN., Abbey, H., & Pless, B. 1990. "Evaluation of an Infant Car Seat Program in a Low-Income Community". AJDC. Vol 144, pp 74-78.

Ref # 1985/2 Sackett, D.L., M., Haynes, R.B. & Tugwell, P. (1985) Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston, Little Brown & Co.

Ref # 1992/2 Sacks, JJ., Brantley, MD., Holmgreen, MS. & Rochat, RW. 1992. "Evaluation of an Intervention to Reduce Playground Hazards in Atlanta Child-Care Centres". American Journal of Public Health. Vol 82(3), pp 429-431.

Ref # 1996/8 Schick, A. (1996) "The spirit of reform: Managing the NZ State Sector in a time of change" Wellington, New Zealand, State Services Commission. http://www.govt.nz.ssc/schickrp.pdf

Ref # 1996/6 Sechrest, L. Stewart, M. Stickle, T.R. & Sidani, S. (1996) Effective and Persuasive Case Studies. Cambridge, Mass., Human Service Research Institute.

Ref # 1994/22 Selby, J.V. (1994) "Case-control evaluations of treatment and program efficacy" in Armenian, H.K. (ed). Applications of the Case-Control Method. Epidemiologic Reviews vol. 16, no. 1, pp. 90 - 101. 1994/22

Ref # 1994/20 Silagy, CA. & Jewell, D. 1994. "Review of 39 Years of Randomized Controlled Trials in the British Journal of General Practice". British Journal of General Practice. Vol 44, pp 359-363.

Ref # 1988/2 Smith, KJ. & Everly, Jr., GS. 1988. "Problems in the Evaluation of Occupational Health Promotion Programs: A Case Analysis". American Journal of Health Promotion. Vol 3(1), pp 43-51.

Ref # 1949/1 Solomon, R.L. (1949) "An extension of control group design" Psychology Bulletin, vol. 46, pp. 137 - 150.

Ref # 1992/5 Steckler, A., Mcleroy, KR., Goodman., RM., Bird, ST. & McCormick, L. 1992. "Toward Integrating Qualitative and Quantitative Methods: An Introduction". Health Education Quarterly. Vol 19(1), pp 1-8.

Ref # 1994/27 Van Beurden, E. (1994) "HOOPS (Health Outcome Oriented Problem Segmentation): A graphical frameworkfor health promotion planning, program development and evaluation" Health Promotion Journal of Australia, vol. 4 (2) pp. 4 - 8.

Ref # 1996/6 Villar, J. & Carroli, MD. 1996. "Theoretical Epidemiology: Methodological Issues of Randomized Control Trials for the Evaluation of Reproductive Health Interventions". Preventitive Medicine. Vol 25, pp 365-375.

Ref # 1986/1 Wan, T. 1986. "Evaluation Research in Long Term Care. Scope and Methodology". Research on Aging. vol 8(4), pp 559-585.

Ref # 1989/2 Webne, S.; Kaplan, B.J.; and Shaw, M. 1989. "Pediatric Burn Prevention: An Evaluation Of The Efficacy Of A Strategy To Reduce Tap Water Temperature In A Population At Risk For Scalds". Journal of Developmental and Behavioural Pediatrics. Vol 10(4), pp 187-191.

Ref # 1993/4 Wiener, RL., Pritchard, C., Frauenhoffer, SM. & Edmonds, M. 1993. "Evaluation of a Drug Free Schools and Community Program: Integration of Qualitative and Quasi-Experimental Methods". Evaluation Review. Vol 17(5), pp 488-503.

Ref # 1993/5 Witte, K., Peterson, TR., Vallabhan, S., Stephenson, MT., Plugge, CD., Givens, VK., Todd, JD., Becktold, MG., Hyde, MK. & Jarrett, R. 1992-93. "Preventing Tractor Related Injuries and Deaths in Rural populations: Using a Persuasive Health Message Framework in Formative Evaluation Research". International Quarterly of Community Health Education. Vol 13(3), pp 219-251.

Ref # 1996/2 Ytterstad, B. 1996. "The Harstad Injury Prevention Study: Community Based Prevention of Fall-Fractures in the Elderly Evaluated by Means of a Hospital Based Injury Recording System in Norway". Journal of Epidemiology and Community Health. Vol 50, pp 551-558.

Ref # 1997/1 Zwerling, C., Daltroy, L.H., Fine, L.J., Johnston, J.J., Melius, J. &. Silverstein, B.A. (1997) "Design and Conduct of Occupational Injury Intervention Studies: A Review of Evaluation Strategies" American Journal of Industrial Medicine vol. 32, pp.164-179