EVALUATING COMPLEX COMMUNITY-BASED HEALTH PROMOTION: ADDRESSING THE CHALLENGES

GWYNETH JOLLEY

South Australian Community Health Research Unit,
Southgate Institute for Health, Society and Equity
School of Medicine
Faculty of Health Sciences
Flinders University

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

May 2013

Contents

Tables

Figures

SUMMARY

This thesis by published work investigates evaluation of community-based health promotion initiatives which use structural or policy approaches rather than focussing on individuals. Empirical research providing evidence of the effectiveness of community-based health promotion is limited.

The thesis consists of a literature review, five papers from my research, and lessons drawn from reflection on my experience as an evaluator of community-based programs. Three of the five papers report on evaluations, including a meta-evaluation of sustainability in a Healthy Cities project. One paper is a review of (mostly) practitioner-actioned evaluations of community health services programs and the other paper reports on arising research leading to the development of evaluation resources.

The research questions are: i) What was the health promotion and evaluation context for my publications and how did this influence my evaluation work? ii) How do my publications reflect evaluation developments prior to 2008, including the role of the evaluator in relation to community-based health promotion initiatives? iii) What are the contemporary challenges in conducting community-based health promotion evaluations? iv) What are the overall lessons from the evaluation practice presented in my publications and how do they inform new approaches to evaluation of community-based health promotion initiatives?

The thesis argues that contested understandings of health promotion and the dominance of a positivist research paradigm present challenges to effective evaluation of community-based health promotion initiatives. Although evaluation

theory has evolved to include interpretive approaches, mainstream evaluation practice still has to contend with demands for a linear, objective scientific approach that does not sit well with community-based health promotion. My evaluation work and the arising publications illustrate the tensions and compromises in taking a more interpretive approach. This thesis contends that, as evaluation has come to be accepted as a more values-based enterprise, health promotion evaluation should reflect principles of participation, empowerment and equity.

Community-based health promotion initiatives are often complex interventions in complex settings and this presents evaluation challenges. These include flexible goals, diverse settings and participants, interaction between stakeholders and dynamic, non-linear programs. The developmental nature of many health promotion programs means that evaluations are context-contingent and this limits transferability of findings.

The thesis concludes that mainstream approaches to evaluation are not able to cope well with the complexity of community-based health promotion and that complexity theory shows promise in addressing evaluation challenges. I classify my studies into complicated or complex domains by examining the extent and diversity of components, stakeholders and interactions and consider how the evaluations might have benefited from use of complexity theory.

Building on insights from my publications, complexity and developmental evaluation, I present a conceptual model of my thinking about planning and evaluation processes. This model brings together program theory and developmental evaluation and may assist evaluation of complex interventions by supporting reflexive practice that can accommodate the adaptive and interactive nature of

community interventions. The thesis argues that ideas from complexity can help to build cumulative evidence in order to identify the foundation principles of effectiveness that can be transferred to a new situation.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain material previously published or written by another person except where due reference is made in the text.

Signed ……  . Date ……. 21st May 2013

ACKNOWLEDGEMENTS

Firstly, I would like to thank the many participants in the studies contributing to this thesis, including health promotion program managers, practitioners and program recipients. You have taught me much over the years and shown such dedication to the task of improving population health and wellbeing.

I sincerely thank the co-authors of my publications: Fran Baum, Patricia Carter, John Coveney, Denise Fry, Richard Hicks, Catherine Hurley, Paul Laris, Steve Parker, Kate Saint and Alison Smith for their contributions to the papers and for allowing me to use this work as part of my PhD thesis. It has been a pleasure to work with you and to share experiences.

I offer my deepest gratitude to my principal supervisor, Professor Fran Baum, co-supervisors Professor Lynn Kemp and Dr Angela Lawless, for their time, support and academic guidance. In particular, Fran has been a source of inspiration since I started working at the SA Community Health Research Unit in 1994, with her wealth of knowledge and experience in public health research combined with a strong sense of social justice. She has provided invaluable guidance on the content and structure of the thesis and particularly on the health promotion aspects. Lynn has been supportive and encouraging throughout, particularly for the evaluative components, and helped to structure the thesis. I have worked closely with Angela for many years and she has always been willing to share ideas and resources, and to act as sounding-board for developing thoughts.

I also owe thanks to all my colleagues at the SA Community Health Research Unit for their interest and support during my period of study, in particular, Dr. Toby Freeman for invaluable assistance with Endnote, and Carly Gowers and Michael Cox

for help in formatting and layout of diagrams. I would also like to thank Patricia Lamb and Helen Scherer for administrative support.

Lastly, thanks to my family, particularly my daughters, Heather and Helen, for their faith in me. I hope that, in a small way, I have helped to inspire their love of learning and that this will pass on to my wonderful grandchildren Bridie, Jamie and Phoebe. Finally, to my parents, Harry (also known as Jack!) and Edna, thank you for teaching and supporting me, through good times and bad.

And to friends and colleagues everywhere who have taken an interest and wished me well, thank you.

ABBREVIATIONS

SACHRU …… South Australian Community Health Research Unit

WHO …….. World Health Organization

CHAPTER ONE: INTRODUCTION

1.1 Introduction

This thesis by published work investigates evaluation theory and practice and relates this to the evaluation of complex community-based health promotion programs. The thesis draws on Australian and international literature, five papers from my previous evaluation research, and lessons from reflection on my experience as an evaluator of complex community-based programs over 15 years. The aim of the thesis is to describe, assess and contribute to addressing the theoretical and practical dilemmas arising from the evaluation research presented in my publications. This will contribute to further advancement of evaluation theory and practice in order to increase researcher and practitioner capacity to undertake meaningful evaluation in community-based health promotion settings.

The thesis undertakes an examination of historical and contemporary thinking on the evaluation of complex community-based health promotion initiatives such as Healthy Cities, other healthy settings initiatives and health promotion programs that use structural or policy approaches rather than an individual behaviour focus. Evaluation of these programs is often limited to process evaluation and, while this is important, there is an increasing demand for measurable outcomes. Making a causal link between a health promotion program and any observed changes is difficult in a community setting so using theoretical explanations of how and why potential outcomes could be attributable to an intervention is a valuable approach. This move to theory-based evaluation approaches represents a major shift from methods-based

evaluation and the subsequent implications for health promotion evaluation are the main focus of this thesis.

The thesis reviews academic literature to trace the development of evaluation approaches and different understandings of health promotion principles and practice. My evaluation work and the arising publications illustrate the tensions and compromises in taking a more interpretive approach to evaluation while working within a still predominantly positivist paradigm.

The thesis then uses insights from more recent complexity theory combined with developmental evaluation, to examine the role that these ideas might contribute to overcoming some of the challenges in undertaking evaluation in complex community-based health promotion environments.

1.2 Context and significance

Empirical research providing evidence of effectiveness of community-based health promotion initiatives is limited. For example, despite recognition of the need to make evaluation integral to Healthy Cities projects since their inception (Baum, 2000; Baum & Cooke, 1992; Boonekamp, Colomer, Tomas, & Nunez, 1999; Burton, 1999; de Leeuw, 1999; Dooris, 1999; World Health Organization, 2000) in practice, this rarely happens. There is a need to build the evidence base for health promotion programs (Judge & Bauld, 2001) and this would strengthen the case for investment in health promotion initiatives such as Healthy Cities, increase credibility and development of a sound theoretical framework, improve implementation and accountability, and build a resource of knowledge (Baum, 2002; Burton, 1999; de Leeuw, 1999; World Health Organization, 1999). The lack of rigorous evaluation of community-based health promotion initiatives is linked to the methodological

challenges this presents. Challenges comprise the following main issues: i) complexity of community-based health promotion initiatives and their settings, ii) using appropriate research approaches and iii) attribution and demonstrating causality.

1.2.1 Complexity

Healthy Cities and similar initiatives are an example of the settings approach to health promotion. As such they are dynamic, complex systems with each setting functioning as an open system in exchange with the wider environment and other settings (Dooris, 2005). Community-based health promotion initiatives tend to have long-term goals that may change over time, multiple actions and expected outcomes at multiple levels and are active in local contexts that differ from setting to setting (Baum, 2002; Judge & Mackenzie, 2002). Further, many of the social processes underpinning action, such as empowerment and community participation, are poorly theorised or are contested in meaning (Baum, 2003; Evans, Hall, Jones, & Neiman, 2007). Recent developments in complexity theory and its application to health promotion evaluation are explored in Chapter Four.

1.2.2 Research approaches

Since the 1970s there has been heightened interest in rigorous examination of the effectiveness of medical interventions and evidence-based medicine, with the development of a hierarchy of evidence with the randomised controlled trial at the top. Braveman and colleagues (2011) note that medicine seems to be unique in the primacy given to randomised controlled trials. However, this is problematic for community-based health promotion where randomisation into experimental and control groups, identical except for exposure to the intervention, is unrealistic (Tones & Green, 2004). While there are some proponents of the use of community

controlled trials (see, for example, Oakley, 1998, 2005), in general setting up control communities and keeping them uncontaminated by the intervention is not practical and, since the initiative is likely to be developmental, it is not possible to predict the exact nature of the intervention or the expected outcomes in advance (Baum, 2002). This means that the notion of the superiority of the randomised controlled trial and other experimental methods has been challenged and a mix of quantitative and qualitative methods to suit the specific evaluation question is proposed by many commentators (see, for example, Baum, 1995; Judge & Bauld, 2001; Nutbeam, 1999). Green and Kreuter (1999) argue that community-based, ecological approaches to health promotion are not so thoroughly evaluated as clinical interventions because the units of analysis – family, community, physical and social environments – do not lend themselves to random assignment to experiment / control groups, nor to manipulation of independent variables. The use of different evaluation approaches and methods is discussed in Chapter Two.

1.2.3 Attribution and causality

The complexity of community-based health promotion initiatives and the use of non-experimental methods mean that a linear model of causality cannot be established with any certainty. The long time frame required for achieving outcomes from many community-based health promotion initiatives adds to the problems of causality and attribution of effect. According to Judge and Bauld (2001) health promotion programs are rarely designed with evaluation in mind, they lack clear documentation of planning and implementation and often have vague goals. For example, Health Action Zones in the United Kingdom, tasked with tackling health inequalities, were required to set out clear plans of how they would achieve social change in the long-term and how they would deliver on specified targets in the short-term. Judge and

Bauld (2001) note that, although strong on identifying problems and long-term goals, the Health Action Zone plans were seldom able to articulate the intervention steps between problems and achievement of goals. Theory-driven approaches to evaluation attempt to overcome attribution and causality questions and this is explored in Chapter Two.

A further challenge to evaluation of community-based health promotion is that research grant bodies favour linear, defined approaches to research and evaluation (Israel, Schultz, Parker, & Becker, 1998; Kavanagh, Daly, & Jolley, 2002) so resources and funding for evaluation of these initiatives has been limited (Evans, et al., 2007). This thesis contributes to methodological development of evaluation approaches that in turn should encourage more rigorous assessment of community-based health promotion and assist research funding bodies to consider how they can better fund flexible and complex research methodologies for evaluation of health promotion interventions.

1.3 Aim and research questions

This thesis investigates theory and practice relating to evaluation of complex community-based health promotion programs. The aim of the thesis is to describe, assess and contribute to addressing the theoretical and practical dilemmas arising from the evaluation research presented in my publications. This will contribute to further advancement of evaluation theory and practice in order to increase researcher and practitioner capacity to undertake meaningful evaluation in community-based health promotion settings.

The following research questions reflect the content of the publications and form the framework for this thesis:

1. What was the historical health promotion and evaluation context for my publications and how did this influence my evaluation work?

2. How do my publications reflect evaluation developments prior to 2008, including the changing role of the evaluator in relation to community-based health promotion initiatives?

3. What are the contemporary theoretical, methodological and practical challenges in conducting community-based health promotion evaluations?

4. What are the overall lessons from the evaluation practice presented in my publications and how do they inform new approaches to evaluation of community-based health promotion initiatives?

1.4 Scope

The thesis takes as its theme the evaluation of complex community-based health promotion initiatives. In Australia, generally, these initiatives are funded and implemented by the Federal or State governments, but some are implemented through non-government organisations with government funding.

Many of the terms used in this thesis, such as ‘community’ and ‘health promotion’ have contested meanings. In order to make my approach clear and to scope the thesis appropriately, a discussion of terms is presented. While there are many other understandings of health promotion and evaluation terminology those chosen here are consistent with the underlying epistemology of the thesis. These definitions have assisted with delineating the literature review and making sense of the arguments presented in the literature.

Health promotion history, strategies and principles are described in Chapter Two. For the purpose of this thesis the definition from the Ottawa Charter is accepted:

> Health promotion is the process of enabling people to increase control over and to improve their health (World Health Organization, 1986a).

This definition makes clear the empowerment principle of health promotion; that is health promotion should be participatory and work to empower people in the organisation, care and management of their health. The WHO definition reflects a view of health promotion acting to increase positive health rather than a focus on the prevention of disease. This salutogenic (Antonovsky, 1996), or health enhancing, approach to health is in contrast to the more medical concept of the pathogenic or disease-causing approach.

Health promotion programs and initiatives aim to promote health in individuals or populations through changing attitudes, beliefs, knowledge or behaviour, or they address structural change, such as healthy public policy, legislation and other action to increase supportive environments for health. For my purpose, treatment, screening and illness prevention activity such as immunisation are excluded from my delineation of health promotion as used in this thesis.

While health promotion interventions can fall along a continuum from individual, family, community and structural (Baum, 2002; Labonte, 1992), my main focus is on evaluation of those initiatives that work at community or population level, and that recognise the impact of the social and environmental determinants of health and how these influence the way people respond to programs and what outcomes are achieved. The programs that are the subject of three of my publications fit this approach while two publications cover work with individuals in addition to communities.

Communities are often described as based on geographical or common interest groups (Fry & Baum, 1992). A geographical community might be defined as a given

local government area or the catchment area for a health service. Communities of interest might be defined by a common social activity or a shared cultural identity. Of course, it would be naïve to believe that such a defined community is homogenous and many writers on the notion of community (for example, Baum, 2002; Jewkes & Murcott, 1998; Petersen & Lupton, 1996; Young, 1990) have pointed out that power struggles and differences are inevitable, even within what might be considered a tight-knit group. This raises the issue of power and control when conducting evaluations of programs and the role of the evaluator and other stakeholders in addressing this. I return to these issues in Chapters Three and Four.

Health promotion activity can occur in a wide range of settings. Neighbourhoods, schools, workplaces, primary health care services and shopping centres are examples of community settings. This thesis is concerned with evaluation of health promotion that is set within a geographical community setting such as a city, or within a community of interest such as a school, workplace or community centre. Health promotion may also occur in residential facilities, such as hospitals and aged care facilities; however, these fall outside the scope of this thesis.

1.5 Publications contributing to the thesis

The following publications form the basis of this PhD by published work. The publications span eight years of evaluation practice and illustrate my development as an evaluator over this time. All the evaluations are of 'real life' programs and draw together theory and practice.

1. Smith, A., Coveney, J., Carter, P., Jolley, G. and Laris, P. (2004) The Eat Well SA project: an evaluation-based case study in building capacity for promoting healthy eating. *Health Promotion International* 19 (3) 327-334.

2. Baum, F. Jolley, G. Hicks, R. Saint, K. & Parker, S. (2006) What makes for sustainable Healthy Cities initiatives? - a review of the evidence from Noarlunga after 18 years, *Health Promotion International*, 21 (4) 259-265.

3. Jolley, G. Lawless, A. Baum, F. Hurley, C. and Fry, D. (2007) Building an evidence base for community health: a review of the quality of program evaluations. *Australian Health Review*, 31 (4) 603-610.

4. Jolley G. Lawless A and Hurly C. (2008) Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion. *Health Promotion Journal of Australia*, 19 (2) 152-157.

5. Jolley G. (2008) Evaluation of an action research project in workforce development and organisational change: Healthy Ageing-Nutrition *Evaluation Journal of Australasia*, 8 (1) 11-19.

1.6 Background and approach

This thesis by published work has its origins in the research and evaluation reported in the five papers which make up the published contribution to the thesis. Since the work covers some eight years and dates back to an evaluation undertaken in 2001, my first task was to re-visit the papers with a view to identifying a common thread that would inform the theme of the thesis. The papers centre on various aspects of the evaluation of community-based health promotion programs. Three describe and report on program evaluations, including a meta-evaluation of a local Healthy Cities

project. One paper is a review of (mostly) practitioner-actioned evaluations of programs situated in community health services and one reports on research arising from that review. Following my re-reading of these papers, the common thread identified centred on the challenges and issues in evaluation of complex, community-based initiatives that are developmental in design. Initial research questions, focussing on synergies and developments within health promotion and evaluation disciplines over time, were developed and refined following a review of the literature.

The evaluations I have conducted and the resulting publications presented as part of this thesis have been set in practice contexts. They were all undertaken during my employment at the South Australian Community Health Research Unit (SACHRU) as part of the work of that agency. SACHRU is part-funded by the South Australian health department (SA Health) with one role being to provide research and evaluation advice and support to South Australian state-funded primary health care services. Other revenue comes from competitive research grants, consultancy work and workforce development services. The work of SACHRU, and my role within it, has to balance a requirement to bring in funding, with a value-base that guides the type of work undertaken. This value-base includes a focus on issues of equity and the social determinants of health and wellbeing, using a participatory and collaborative approach wherever possible, and undertaking research and evaluation that is independent, rigorous and disseminated widely and in appropriate ways (South Australian Community Health Research Unit, 2012). The influence of this work context on the nature of the research and evaluation reported in my publications is considered later, in Chapter Three.

My experience in evaluating community-based initiatives suggests that the evaluator needs to 'get their hands dirty' rather than remain an 'objective' outsider. The evaluation task usually starts by needing to clarify the underlying understandings (or theories) of how and why the program is expected to achieve results. This is done by engaging with the various stakeholders and becoming, in effect, a partner to the program. In my experience, this assists in understanding the complex dynamics at play and ensuring that the evaluation is of practical use to practitioners and decision makers. My approach is that the theory of the program should inform evaluation practice and that out of practice new program theory can be grown. The legitimacy of this approach is considered in the literature review (Chapter Two) and is also the subject of further reflection and analysis in Chapter Three.

My publications considered in chronological order offer insight to the progression of evaluation theory and approaches as they relate to health promotion. The writing of this thesis has required me to reflect on the lessons from my own work and the issues and challenges in evaluating community-based initiatives.

My initial inclinations and training were towards a positivist and scientific view of the world. Although I was, as a child, an avid reader, my first love at secondary school was in science and mathematics. I took pleasure in using the formal language of science reports and drawing the precise diagrams accompanying them. I also enjoyed the challenge of systematically working through a maths problem and the certainty of knowing there was one correct solution. My early working life continued on this trajectory, working in a science education laboratory while studying for my B.Sc. and then teaching science and maths at secondary level for some ten years. My evaluation experience at this stage was limited to assessing and marking students'

work. However, the move to Australia and a number of personal events brought about changes to my life that, while challenging, opened up opportunities to re-consider my world view. In 1994, I started working at the South Australian Community Health Research Unit and was immediately confronted with a whole new (to me) paradigm of qualitative enquiry. While studying for the Master of Science (Primary Health Care) at Flinders University I was fortunate to have the late Michael Crotty as a lecturer. Michael's skill and passion for qualitative research was inspirational and his book *The Foundations of Social Research* (Crotty, 1998) is a continuing source of knowledge. Professor Fran Baum was also teaching on the Masters course and I found my eyes opened to the importance of the social and economic determinants of health, health equity and ethical public health research.

These two major shifts in my thinking have, of course, influenced my work as an evaluator. Completion of my Master's thesis on the use of performance indicators for community health services (Jolley, 2003) convinced me of the inappropriateness of relying solely on quantitative indicators to assess these services. By the time of the evaluation that is described in my first publication for this thesis, I was gaining expertise in using both quantitative and qualitative approaches to data collection and analysis and was developing skills in designing and using program logic models and related program theory.

Now, at the time of writing this thesis, I am what Chen (2005) describes as a 'contingency' evaluator; that is choosing approaches and methods that best fit the evaluation situation. The proviso is that the approach complements my skills in order for me to perform adequately, and the implicit or explicit value system of the intervention being evaluated must be compatible with my own and that of my

organisation. As noted earlier, my standpoint on health promotion is grounded in the Ottawa Charter and a valuing of the underpinning principles of empowerment, participation, equity and sustainability as described by a number of commentators (see, for example, Keleher, 2007; Tones & Green, 2004). It follows that my approach in this thesis is centred on community-based health promotion initiatives that recognise these underlying principles, and on evaluation design that also reflect these.

Lessons from reflecting on my evaluation work over the time period of my presented publications form part of this thesis and are presented in Chapter Three where I draw together the findings from the literature review and my publications.

1.7 Structure of thesis

The thesis is presented in five chapters. Chapter One, this chapter, describes the background to the thesis, the aims and my approach. It presents a scoping framework for the thesis and outlines my position as an evaluator at the start of the journey that my publications describe.

Chapter Two is the literature review, structured into two sections: i) a review of different understandings of health promotion and how these impact on evaluation and ii) a developmental history of evaluation theory and approaches and the implications of these for health promotion evaluation. The literature review sets the background for my publications and the context in which my evaluation work was undertaken. It provides a framework for drawing together the lessons and implications from the publications.

Chapter Three contains abstracts and further analysis of the publications submitted as

part of this thesis (copies of the full papers are provided in the Appendix). For each publication, there is a description of the study and its context, my theoretical position at that time, and the methodological and practical challenges.

Chapter Four looks at the complexity literature that has emerged since the evaluation work that forms part of this thesis and discusses more recent notions of complexity theory and developmental evaluation that post-date my publications. In this chapter I consider the contribution that complexity theory might make to health promotion evaluation and I propose a conceptual model, drawing on my published work and complexity theory and developmental evaluation, that illustrates my thinking and may assist the process of planning, implementation and evaluation of community-based health promotion initiatives.

The concluding chapter, Chapter Five, summarises the findings and answers to the research questions arising from my evaluation work and resulting publications. I describe the contribution to knowledge made by this thesis and the limitations of my evaluation research. In conclusion, I argue that more recent approaches to evaluation that include program theory **and** notions from complexity show potential for community-based health promotion but that there are few examples of practical application. The proposed Planning, Implementation and Evaluation conceptual model illustrates how these ideas can be brought together to address some of the challenges in community-based health promotion evaluation.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

This thesis is concerned with evaluation of community-based health promotion. Therefore I have explored the literature in these two fields in order to examine different understandings of health promotion and the implications of these for evaluation, and to examine different approaches to evaluation and their synergy with community-based health promotion.

The purpose of the literature review is to describe and critique the health promotion context (and specifically community-based health promotion) and to trace the historical development of evaluation that formed the background for my publications. This provides the evaluation context at the time for my work and the arising publications, and forms the foundation for the description and analysis of my work as an evaluator of community-based health promotion. I also consider, in Chapter Four, more recent evaluation literature that post-dates my published work in order to examine some contemporary approaches to addressing the challenges of evaluating community-based initiatives.

In this chapter, following a methods section, the review is in two main parts. The first examines the development, principles, values and practice of health promotion, particularly as practiced in community-based settings. The second part details the development of evaluation and evaluation theory and its relevance to community-based health promotion initiatives. The aim is to trace and critically examine developments in the fields of health promotion and evaluation over the period covered by my published evaluation work in order to investigate the context for both at the time, and how this influenced my work. The literature review illustrates the

challenges and issues for community-based evaluation and how evaluation developments have attempted to address these.

The literature review is narrative in that its purpose is to generate understanding and reflect a process of discovery, rather than accumulate knowledge (Bryman, 2012). According to Bryman (2012), a narrative literature review is suitable for qualitative research based in an interpretive epistemology. This approach accords with my evaluation practice and the focus of my publications which mainly use a qualitative methodology as most appropriate for the type of evaluation work conducted.

2.2 Literature search strategy

As a result of my (and colleagues') research and evaluation of health promotion and community health programs, I already had established a bibliographic database with references gathered over time. This database contained citations from my publications, references specific to Healthy Cities and seminal evaluation and health promotion literature. Thus, the literature review began with works already known to me and this was supplemented by following up citations from these works and by recommendations from colleagues. Publisher and library alerts were used to access information about relevant new works. Following further reading and early drafts of the literature review the research questions were made more specific. As gaps around specific topics were identified I used database searches, primarily Medline, PubMed, CINAHL, and Google Scholar. Search terms 'health promotion', 'community-based health promotion', 'healthy settings', 'Healthy Cities' and 'evaluation' were used in combinations. To identify the complexity literature reviewed in Chapter Four, the terms 'complexity theory'; 'evaluation'; and 'health promotion' were combined to identify works describing theory or practice in this area. Quality was assessed

through relevance to my research questions rather than explicit quality criteria as used in systematic reviews. This means there may be gaps in terms of particular studies but the review is comprehensive in identifying key evaluation theorists and approaches.

2.3 Health promotion

This section outlines a brief history of health promotion and its development. Different understandings of health promotion are discussed and some underlying principles and values established. The principles and practice of community-based health promotion are described using the settings approach of Healthy Cities as an example. Finally, the impact of the principles/values and practice of community-based health promotion on evaluation approaches is described. Specific issues relating to the evaluation of these types of initiatives are discussed.

2.3.1 Developmental history of health promotion

Modern understandings of health promotion date from the 1970s. Health promotion has its roots in health education, originally concerned with increasing knowledge about environmental influences on health, such as supply of clean water. By the mid-twentieth century, health education was focussed on using information dissemination to persuade individual behaviour change and according to Green and Tones (2010) this took two approaches: preventive health education based on psychological theory, and a more enabling approach based on educational theory. McQueen (2007) suggests that there are geographical differences, with health promotion in the United States rooted in health education and psychology whereas in Europe and Canada health promotion developed from a more socio-economic and political base. In Canada, the Lalonde report (Lalonde, 1974) returned the focus to the social and

economic factors influencing health by recognising the impact on health of environmental and lifestyle factors, in addition to the traditional focus on the biological determinants of health (Talbot & Verrinder, 2010). The Lalonde report was also the first government report to publicly acknowledge that the health care system was not the most important factor in determining health status (Hancock, 1986). However, Ashton and colleagues (1986) argue that, following widespread global adoption of a particular interpretation of the Lalonde report, health promotion tended to victim blaming lifestyle approaches and acceptance that people were powerless to change environmental factors. Later commentators claim that the Lalonde report laid the foundation for new ways of thinking about health promotion and that the report argued for social change through collective action rather than the notion of people being powerless to change their environments (Labonte & Penfold, 1981).

In the late 1970s, it was becoming apparent that provision of medical care had reached a point of diminishing returns; with increased health costs (Jirojwong & Liamputtong, 2009a) and limited effects on population health (Wass, 2000). There was also growing awareness of the inequity of high cost care available for a minority of people and a lack of basic health care for many, in both developed and developing countries (Wass, 2000). This recognition of the limitations of medical interventions, along with the rise in awareness of the importance of lifestyle as a driver of, mostly chronic, disease, contributed to the 1978 Declaration of Alma Ata with the aim of *Health for All by the year 2000* and the understanding that health promotion, along with prevention, curative and rehabilitation services, is an essential component of health systems. The Alma Ata Declaration, with an emphasis on equity, social justice and empowerment, called for a major change in how health services operate and the

political will of governments to carry through the change (Wass, 2000). The role for health promotion was formalised in the Declaration, which stressed the part to be played by health promotion in achieving the goal of *Health for All*.

Meanwhile, in Australia, a period of radical social change resulted in the Community Health Program in 1973 to complement Medibank, the then new universal public health insurance scheme. Under the Community Health Program multi-disciplinary community health centres were established with responsibility for the health of a given area (Owen & Lennie, 1992) and emphasising illness prevention. Raftery (1995) argues that, while relatively short-lived, the Community Health Program laid the foundations for a new approach to public health in Australia and so helped in the local acceptance of the WHO *Health for All* program.

Health for All emphasised primary health care as the main way to address health problems (initially in developing countries) and recognised health promotion and disease prevention as important strategies within this. Other elements of *Health for All* included recognition of the importance of global cooperation and peace and the broader social and economic determinants of health, the achievement of equity in health status, participation by people in planning, organisation and control of health care, and involving all sectors in the promotion of health (Baum, 2002).

This broad agenda for primary health care linked health improvement to social and economic development in developing countries and was taken up by many nations (Wass, 2000). However, two streams of primary health care soon emerged. The intent of *Health for All* was for comprehensive primary health care concerned with social justice, equity, community control and social change, focusing on ‘conditions that generate health and ill-health’ (Wass, 2000 p12). A more selective approach

(Walsh & Warren, 1979), giving priority to medical interventions over social action to improve health, and focussing on individual disease, quickly took hold. The selective approach maintains professional control and takes the view that medical care is the major determinant of health (Wass, 2000). The medical model, defining health as the 'absence of disease', leads to a focus on specific diseases and on lifestyle education interventions, such as media campaigns urging individuals to reduce risk factors, for example, smoking, alcoholism and obesity, associated with poor health (Baum, 2002). Wass (2000) notes that two challenges to comprehensive primary health care hindered its acceptance; firstly, the notion that primary health care is only relevant in developing countries, and secondly, medicine and the medical industry which stand to lose power in a broader implementation of comprehensive primary health care.

Green and Tones (2010) argue that the history of health promotion similarly has featured a struggle to distance itself from association with the medical model of health. These authors argue that applying the medical model to health promotion leads to a focus on illness prevention and a preventive model that is concerned with risk and with individual responsibility to reduce exposure to risky behaviour and risky environments, and that this is incompatible with health promotion values of equity and empowerment. Instead, Green and Tones propose a 'new health education' that provides opportunities for learning and thus to increased control over personal health and health in the community.

The 1980s saw debates in the health promotion sector over 'education' versus 'behavioural and environmental change' as it became apparent that education alone was insufficient to bring about change in behaviour related to complex socially

embedded lifestyles (Grembowski, 2001). Green and Tones (2010) argue that the traditional form of health education, with its individual focus, is based on a preventive medical model of health and tends to victim blaming. There are also equity and social justice concerns that education and behaviour change strategies are less effective for lower socio-economic populations (Grembowski, 2001).

The shift from purely educational to behavioural and environmental understandings of health promotion was followed in 1986 by the Ottawa Charter for Health Promotion which states the prerequisites of health as including peace, a stable eco-system, social justice and equity (World Health Organization, 1986a). Health promotion is thus based on similar principles to primary health care (Talbot & Verrinder, 2010; Wass, 2000). The Ottawa Charter identifies three basic strategies for health promotion: advocacy for health to create the essential conditions for health; enabling all people to achieve their full health potential; and mediating between the different interests in society in the pursuit of health. Thus, the Ottawa Charter supports a shift of focus from proximal to more distal risk factor risk conditions (Grembowski, 2001).

The basic strategies of the Ottawa Charter are supported by five priority action areas:

- Build healthy public policy
- Create supportive environments for health
- Strengthen community action for health
- Develop personal skills, and
- Re-orient health services (World Health Organization, 1986a)

Thus, action for health promotion is required in multiple areas: with individuals and

groups; in policy making; in social environments and in health services themselves. This ability of the Ottawa Charter to integrate systemic approaches to health with behaviour and lifestyle factors contributes to its enduring strength (Baum, 2002). In 2011, the 25th anniversary of the Charter prompted reflection on its current relevance with most commentators agreeing that the strategies of enabling, mediating and advocacy remain relevant despite the enormous changes in social and environmental context (Kökény, 2011; Saan & Wise, 2011).

The definition of health promotion used in the Ottawa Charter is 'Health promotion is the process of enabling people to increase control over, and to improve their health' and this emphasises the empowering nature of health promotion rather than a focus solely on education for individual behaviour change. The WHO Glossary (World Health Organization, 1998) goes on to elaborate further:

> Health promotion represents a comprehensive social and political process, it not only embraces actions directed at strengthening the skills and capabilities of individuals, but also action directed towards changing social, environmental and economic conditions so as to alleviate their impact on public and individual health. Health promotion is the process of enabling people to increase control over the determinants of health and thereby improve their health.
>
> Participation is essential to sustain health promotion action. (World Health Organization, 1998 p1-2)

This definition strengthens the notion of health promotion as a social and political activity and notes the importance of addressing the social, environmental and economic determinants of health while also recognising the importance of personal skills and capabilities. It also confirms the importance of citizen participation in health and health decision making.

The Ottawa Charter confirmed the need for multiple strategies including personal skills in health, supportive environments for health and healthy public policy. Since

the Ottawa Charter in 1986, international health promotion conferences have been held regularly, each with its own themes and areas for development. The Adelaide, Australia conference, 1988, took public policy and policies to reduce inequities as its themes, then, in 1991, the Sundsvall, Canada conference focused on supportive environments, strengthening advocacy and empowerment (Catford, 2011; Wass, 2000). In 1997, the Jakarta Declaration (World Health Organization, 1997) added another five priorities for health promotion that confirm the need for governments to invest in health promotion and for all organisations and sectors to work together to advance health. In 2000, the 5th International Conference in Mexico focussed on equity and health determinants (Talbot & Verrinder, 2010), while in 2005 the Bangkok Charter (World Health Organization, 2005) continued the work on equity by discussing the need to tackle threats to health from global development, the fair distribution of resources, rights to health and equity of access to health care, and reconciliation with indigenous peoples. Kenya hosted the 2009 conference which called for urgent action to mainstream health promotion, to strengthen the workforce and leadership, increase participation and empowerment processes and build and apply knowledge (Catford, 2011).

In reflecting on achievements in the 25 years since the Ottawa Charter a number of commentators suggest that there has been uneven progress. Petterson (2011) for example, calls for the WHO to take more responsibility for implementation of outcomes from the global conferences. Hancock (2011) describes the progress of health promotion in Canada as 'unfulfilled promise' citing a reframing to 'population health' with its less political approach and budget cuts as the key reasons.

As with the debates over the direction for primary health care, health promotion is

also a contested term with practice ranging from individual health education and mass marketing of health promotion messages, to support for community action and advocacy for policy and system change. This range of strategies is discussed in the health promotion practice section below. Whatever the stated intention, there seems to be a continual trend for health promotion to focus on individual lifestyle rather than action on the social and cultural, and to emphasise individual behaviours and epidemiology risk factors rather than the social determinants of health (Baum, 2011; Hancock, 2011). Challenges facing implementation of broader health promotion include competition for funding from medicine, the medical dominance of health system policy and practice and the appeal to policy makers of the individual approach and its apparent simplicity (Peersman, 2001).

The different understandings of health promotion have an impact on how health promotion is practiced and provide challenges to evaluation. One way to address these challenges is to consider the various theories and models underpinning health promotion practice and how these have an impact on evaluation. The next section considers these health promotion theories and models and the implications arising for evaluation.

2.3.2 Health promotion theories

Most health promotion theories have been borrowed from social and behavioural sciences and are not highly developed (McQueen, 2007; Nutbeam, Harris, & Wise, 2010). For this reason, Nutbeam and colleagues (2010) suggest that the terms ‘theoretical framework’ or ‘model’ offer a better description as a fully developed theory is often not established. Further, it is important to remember that a theory is a simplified representation of reality and cannot explain all the complexities of individual, social and organisational behaviour (Nutbeam, et al., 2010).

Theories can be useful in the planning, implementing and evaluation stages of a program. For example, theory can provide guidance on whether to focus on individual beliefs or organisational change, on a reasonable standard to benchmark the implementation, and on potential outcomes and ways to measure these (Nutbeam, et al., 2010). Nutbeam and colleagues caution that, since there is no single theory of health promotion, a program may need to be informed by several theories. The task is to find a theory or theories that best fit in helping to explain the link between the program, the problem it is intended to address and the outcomes it seeks to achieve. Moreover, Kickbusch (1997) argues that a single health promotion theory is not necessary since health promotion is a process of social change drawing on theories from many disciplines.

Nutbeam and colleagues (2010) provide a useful overview of health promotion theories and models most commonly used. A summary is shown in Table 2.1 along with an additional column identifying the implications of each theory for health promotion evaluation. As can be seen, behaviour change theories dominate and this may explain why much health promotion activity is directed towards individuals rather than communities or organisational change (Metzler, Amuyunzu-Nyamongo, Mukhopadhyay, & De Salazar, 2007). Hawe and colleagues (2009) argue that most community-based programs rely on aggregating up individual level theory to community level and this results in health promotion interventions with modest or negligible effects becoming the norm, with high costs and marginal benefits.

Table 2.1 Health promotion theories and implications for evaluation (adapted from Nutbeam, et al., 2010)

Approach	Theory or model	Commentary	Implications for health promotion evaluation
Individual behaviour change	Health belief model	Useful for preventative behaviours such as screening. Less useful for complex behaviours such as alcohol/drug use. Does not consider social, environmental and economic determinants	Reasonable evidence that change in individual belief can lead to change in behaviour. Relatively simple to design pre and post testing of beliefs and behaviour change
	Theories of reasoned action and planned behaviour	Assumes rationality. Useful to identify beliefs about causes and what can be done	Limited evidence for success of interventions based on predicted behaviour change
	Stages of change model	Recognises need for range of programs for populations and individuals at different stages of change	Limited evidence of effectiveness. Fails to account for complexity of behaviour change
	Social cognitive theory	Recognises interaction between individual and their environment and concept of self-efficacy. Health practitioner becomes change agent	More comprehensive, recognises need to consider multiple levels of a program and therefore multiple foci for the evaluation
Change in communities and community action for health	Community mobilisation	Encourages involvement of communities and consideration of social determinants of health. Risk is taking a 'deficit' approach rather than building on strengths, and marginalising those least heard.	General lack of evidence in public health for this approach, but can draw on other disciplines. Difficult to implement in practice and therefore difficult to evaluate
	Diffusion of innovation theory	Explains how and why populations respond to new ideas. Does not address structural barriers to adopting change and so risks 'victim blaming'.	Use documented in a wide range of settings. Useful in transferring evidence to practice

Communication strategies to change behaviour	Health literacy model	Goes beyond education to giving people confidence to act. Potential for political action	Measurement of health literacy not well developed and effectiveness not yet tested
	Communication behaviour change theory	Identifies range of issues to be considered for public communication campaigns and shows it is difficult to bring about sustained change by this means alone	Use has demonstrated need for public campaigns to be part of a broader strategy
	Social marketing theory	Health messages targeted at specific populations, to change social norms and behaviour. More complex than social marketing and mutual benefit to individual and society	Research based, easy to understand and do. Reasonable evidence can be effective
Change in organisations and organisational practice	Organisational change theory	Identifies need for staged process involving different organisational levels. Best used for developed programs to be introduced to organisation	Organisational climate culture and capacity needs to be evaluated as these will impact on how program is adopted and sustained'
	Intersector action models	HP needs to work with other sectors to address many health problems. Organisations need to change internally to give capacity for intersectoral work. High investment in relationships may outweigh benefit.	No single model but evidence available on important factors needed for effective action

Development of healthy public policy	Healthy public policy framework	Policy makers, policy influencers, public, media are main stakeholders. Underlies the political nature of policy making and conflicting interests	Multiple stakeholder perspectives, conflicting interests, whose voice is heard all impact on evaluation. Credibility and evidence needed to influence policy
	Evidence base policy making	Evidence used in a variety of ways. Not rational but interactive and political. Assumptions, beliefs, interests and power all influence use of research	Evidence needs to be current and accessible, politically acceptable, and reflect resources and capacity for action. Evaluation needs to identify assumptions etc.
	Health impact assessment	Suggests a way to work across sectors by assessing evidence of health effects of policy on e.g. equity. Makes assumptions more transparent.	Can provide a voice for multiple stakeholders

As can be seen from Table 2.1, these different health promotion theories are supported by varying levels of evidence. In moving from individual behaviour change through to organisational change and development of public policy there are increasing levels of complexity and decreasing levels of substantial theory. This, in turn, increases the complexity for evaluation. Nutbeam and colleagues (2010) conclude that psychosocial theories of individual behaviour change are the least complex and best tested; other theories are less researched and less amenable to experimental design. Peersman (2001) agrees that health promotion theories need to go beyond the psychological and that theories to deal with ecological understandings of behaviour change need to be further developed. In practice, program workers, acting in the contested space of health promotion are likely to draw on a range of theories simultaneously to reflect the different strategies in an intervention: that is individual behaviour change, community mobilisation, organisational change and

development of healthy public policy. Kickbusch (1997) argues that the choice of health promotion strategies should be based on knowledge of how social and behavioural change is best achieved, which, of course, assumes theoretical and evaluative knowledge is available.

Combining the community mobilisation model and healthy public policy framework shown in Table 2.1 above, Tones and Green (2004) propose an 'empowerment model of health promotion' with two major components: health education and public policy. In this model, health education is re-framed as 'critical health education' and its purpose is political and radical, acting as a catalyst for change at policy level, rather than using persuasion to encourage individual behaviour change as traditional health education does. This empowerment model is in contrast to the health education emphasis on modifying lifestyle that assumes people have the capacity to exercise choice in changing their behaviour and which may lead to victim blaming of those who fail to make the right choices (Petersen & Lupton, 1996). Thus, the role of critical health education is empowerment, creating public pressure for change in policy which then leads to more supportive environments for health and more responsive health services (Tones & Green, 2004 p38). Key components of the empowerment model are shown in Figure 2.1.

In this model, education and training leads to individual empowerment and thus to lifestyle changes. Individual empowerment also leads to community empowerment and on to pressure for healthy public policy and healthier environments. Health and medical services are re-framed or re-oriented in order to better meet the health needs of particular populations.

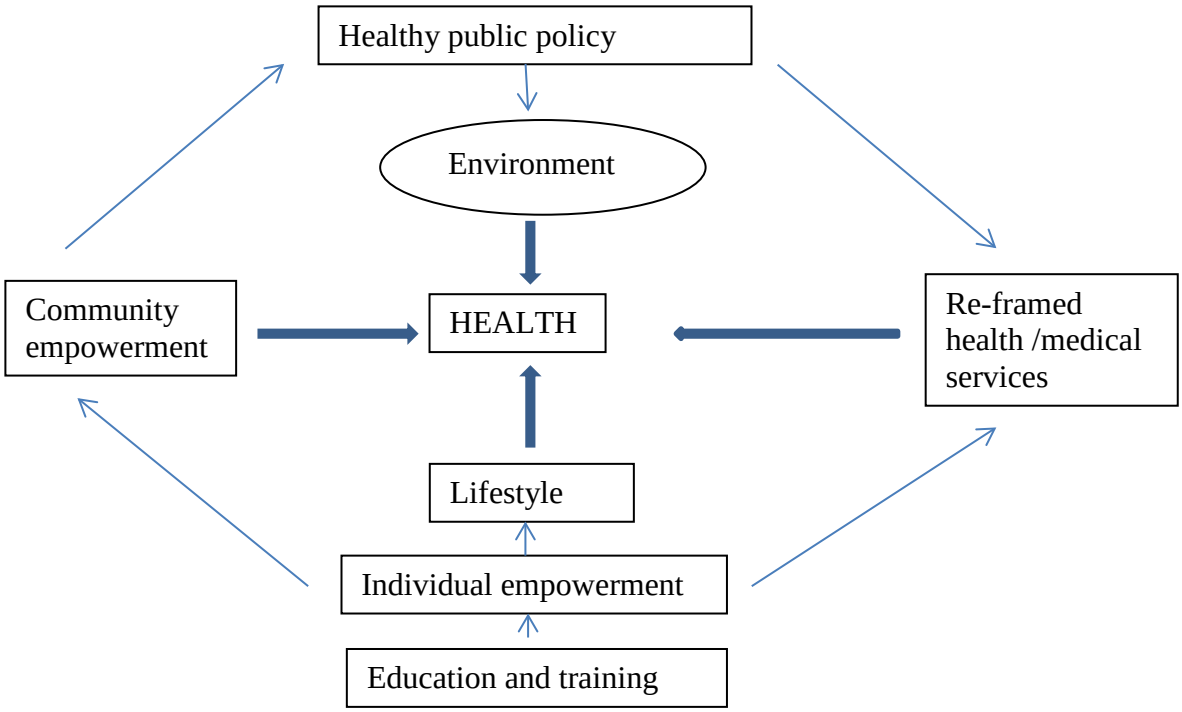

Figure 2.1 Empowerment model of health promotion (simplified from Tones & Green, 2004 p37).

The empowerment model seems to accept as self-evident that empowered individuals and communities are more likely to take part in community action, engage with services and adopt healthy lifestyles. In contrast to this, Peersman (2001 p6-7) states that 'the extent to which empowerment and community participation have improved people's lives or their health remains virtually undocumented'.

Having examined the theoretical underpinnings of health promotion, the next section considers principles and practice.

2.3.3 Principles and practice of health promotion

Health promotion writers have described a set of principles to guide practice that reflect the values of primary health care and the Ottawa Charter. These guiding principles describe ideal health promotion practice as:

- Empowering of individuals and communities to assume more control over factors that affect their health
- Participatory for all concerned, at all stages
- Holistic, to include physical, mental, social and spiritual health
- Intersectoral, with collaboration from sectors other than health
- Equitable, with a concern for equity and social justice
- Sustainable, with changes maintained when funding for an initiative has ended
- Multi-strategy, including a combination of policy development, legislation, organisational change, education, advocacy, community development (Keleher, 2007; Tones & Green, 2004).

For Green and colleagues (2000) the empowerment component of health promotion is critical. They go so far as to say 'The absence of empowering activities should be a signal that an intervention does not fall within the rubric of health promotion' (Green, et al., 2000 p8) and suggest that many health promotion programs labelled as such would not meet this ideal.

Under the banner of these guiding principles, health promotion interventions can operate at a number of levels, for example, individual, group and society (Jirojwong & Liamputtong, 2009b); or individual, community, organisation and public policy (Nutbeam, et al., 2010). Labonte (1992) developed a continuum of intervention strategies (see Figure 2.2) for a 'heart health inequalities' health promotion program in Canada that illustrates the range of health promotion activities that should be considered when planning a program. Labonte considers that these strategies form a continuum from individual empowerment through to political action by coalitions of

community organisations. The strength of this model is that it illustrates the range of approaches likely to be undertaken as a continuum rather than activities based on discrete theories or models. This model is also useful as it identifies potential focus areas for evaluation.

Figure 2.2 Continuum of health promotion strategies (based on Labonte, 1992 p24)

Health promotion programs that operate at multiple levels are most likely to address a broad range of health determinants (Nutbeam, et al., 2010) and a socio-environmental approach to health promotion is more likely to be effective (Labonte, 1992; Wilkinson, 1996). The corresponding increases in complexity, however, mean that evaluation of effectiveness is likely to be more difficult. These challenges to evaluation will be discussed later in the chapter.

Health promotion practice, then, is wide-ranging in terms of theories, strategies and intended participants. Since the focus in this thesis is on community-based health promotion initiatives, the next section describes these initiatives and the settings approach as an example.

2.3.4 Community-based health promotion initiatives and the settings approach

Community-based health promotion comprises activities in communities that draw on the principles of the Ottawa Charter (Baum, 1998). Baum suggests that community-based health promotion should be characterised by the following:

- Recognises different power levels within communities and stakeholders
- Takes a socio-demographic approach using medical, behavioural and community development strategies
- Identifies issues to be addressed by the concerns of community members
- Includes concerns about equity of health outcomes
- Recognises diversity and needs of sub-groups
- Is participatory in planning, implementation and evaluation (Baum, 1998).

In this view, community-based health promotion therefore emphasises empowerment and participation of communities in addressing health issues, uses a range of strategies and is concerned with equity.

The healthy settings approach to community-based health promotion has developed substantially and includes Healthy Cities, hospitals, universities, workplaces and schools. The WHO Healthy Cities project originally established in Europe in 1987 was the first example of a health promoting settings initiative. Settings can be a physical place in space and time where people come together for a specific purpose (for example, a school) or an arena of interaction (for example, a city) (Green, et al., 2000). A settings approach acknowledges the physical, organisational and social contexts in which people live, work and play, as legitimate objects for research (Poland, Frohlich, & Cargo, 2009). In this way, program goals are no longer

focussed on changing specific behaviours in individuals, but move towards creating conditions supportive of health and organisational change for health (Boutilier, Cleverly, & Labonte, 2000). Typically, such initiatives work across sectors, use multiple strategies and work with communities to build engagement and capacity rather than providing services with a top-down approach. That is, lay knowledge is valued and the agenda for priorities, issues and activities is identified by community members rather than health professionals, although Baum (2003) notes that, in practice, issues are often identified by professionals who then go on to try to engage the wider community.

While settings for community-based health promotion initiatives vary, some common principles have been identified by Dooris (2005):

- an ecological model of health, determined by complex interactions between environment, organisation and personal factors, largely outside the control of health services
- salutogenic rather than pathogenic
- works with populations rather than individuals
- holistic rather than single health problems and risk factors
- settings understood as complex dynamic systems, each setting is part of a greater whole
- focus is on bringing about and managing change within the whole organisation or community (Dooris, 2005).

While these common principles have been identified, different approaches are apparent. For example, Tones and Green (2004) describe a number of different types of settings and models but they argue that diversity in the approach allows

responsiveness to local needs and is not of concern as long as consistency with the core principles is maintained. Similarly, Ashton and colleagues (1986) note that from early in its history, the Healthy Cities approach was committed to community participation and intersectoral collaboration and a recognition that there should not be central prescription but that each city would have diverse plans and approaches. This responsiveness to local needs and contexts is a strength of the settings approach but has implications for transferability of evaluation findings as discussed in Chapter Three.

Green and colleagues (2000) argue that the settings approach is critical to health promotion theory because it provides a conceptual boundary for context, and for practice, by defining the people and location for activities of an initiative. According to Boutilier and colleagues (2000) however, the settings approach goes beyond merely providing an intervention in a location, but aims to ensure that the ethos and activities 'are mutually supportive and combine synergistically' to improve health and wellbeing. This frames the setting as a complex environment where people and relationships interact dynamically with health promotion activities. A rather more instrumental description of the settings approach is given by Mullen and colleagues (1994). For these researchers, settings are useful because the interactions and patterns of formal and informal membership and communication channels 'create efficiencies in time and resources for health education programming and offer more access and greater potential for social influence' (Mullen, et al., 1994 p330). So, for these authors, the main benefit is in having a captive audience for dissemination of health messages and implementation of policy, rather than strengthening community capacity and empowerment as advocated in the Ottawa Charter.

Critiques of the settings approach are centred on a number of assumptions about the nature of the setting and power relationships within in it. For example, Green and colleagues (2000) remind us that settings are not homogenous and self-contained but are made up of diverse social relationships with people and organisations entering or leaving the setting. Another assumption, according to Green and colleagues (2000) is that settings are amenable to organisation and policy levers, but the resistance of existing structures and processes is often underestimated (Green, et al., 2000). As these authors point out, there are issues of competing interests within the setting, and gatekeepers to the setting. Further, health promoters usually need to work with gatekeepers or power holders to gain access to a setting, but this may alienate the less powerful. There is also concern that working in a setting can reach those who engage with the setting but miss those on the margins, such as the unemployed (Dooris, 2005; Green, et al., 2000). Baum (2002) raises questions about the extent of genuine participation in settings-based initiatives, given that these programs are typically introduced initially by those in power. These issues bring important context to evaluation because they are likely to influence how an initiative is implemented and the outcomes achieved.

The Healthy Cities movement, as an example of a settings approach in action, has recognised the importance of research and evaluation from its early days. Numerous commentators (see, for example, Baum, 2003; Neiman & Hall, 2007; Poland, 1996a) have iterated the importance of conducting evaluations, developing indicators and establishing causative theories for Healthy Cities initiatives. Despite this, there have been very few published evaluations that go beyond assessment of process issues (de Leeuw & Skovgaard, 2005; Neiman & Hall, 2007; Poland, 1996a). While process evaluation is vital, the call for impact evaluation is growing as funders demand

evidence of the effectiveness of programs in terms of health or social changes (Keleher, MacDougall, & Murphy, 2007). The numerous challenges to establishing evidence to link community-based health promotion programs directly to changes in health status are discussed next.

2.3.5 Health promotion evaluation

The history, principles, theory and practice of health promotion, as described above, all influence evaluation design and conduct. Much modern health promotion continues to be associated with a medical and behavioural paradigm with the main emphasis on illness prevention and the concept of risky behaviours. While alternative perspectives, such as those from a social view of health, have contributed much to health promotion, it continues to be closely linked to health and medical services with a very different understanding of health and illness. Thus the advent of evidence-based medicine has created a surge of interest in 'outcome' evaluation, and this has spilled over to health promotion (Wimbush & Watson, 2000). However, medical outcomes are usually measured in terms of mortality and morbidity and do not encapsulate the positive health outcomes predicted from health promotion interventions. These outcomes may be, for example, increased individual or community levels of empowerment, increased health promoting behaviours or decreased risk behaviours, increased collaboration and action for change by stakeholder groups, implementation of new public policy to support health, and so on. Thus, the most likely outcomes to be observed from health promotion interventions are changes in the conditions that create health and in individual, community and organisational capacity (Baum, 1998). Assessing these types of outcomes requires new approaches to evaluation that can accommodate the complexity of health promotion interventions and the settings in which they are

implemented. The characteristics of community-based health promotion and the implications for evaluation are shown in Table 2.2 and discussed below.

In terms of the setting, Tones and Green (2004) point out that a health promotion setting is culturally constructed, with pre-existing relationships and permeable boundaries. Thus, settings are not discrete, fixed entities but exist as complex systems. This means evaluation of settings-based health promotion initiatives is not conducive to a simple input-output model of cause and effect but rather needs to be able to cope with a complex web of interactions (Tones & Green, 2004). Such an initiative is less amenable to evaluation because it is hard to set parameters and priorities when everything interacts (Green, et al., 2000) and boundaries are unclear (Dooris, 2005). Further, there is great diversity of approach and practice, and variations in settings (Dooris, 2005; South & Woodall, 2012), which implies that evaluation frameworks need to be flexible and diverse in response to this variation and also that transferability of findings is problematic.

Interactive relationships between actors in the setting and between the actors and the environment means evaluation needs to monitor and chart interactions and be alert to how they influence implementation and outcomes. As Poland and colleagues (2009 p505) argue, interventions ‘wither or thrive based on complex interactions between key personalities, circumstances, and coincidences’.

Table 2.2 Community-based health promotion characteristics and evaluation implications

Community-based health promotion characteristics	**Evaluation implications**
Settings context	Context for initiative is important and should not be controlled out
Setting is permeable	Context for initiative and stakeholders are subject to change
People-centred and built on interactive relationships	Initiative is a function of relationships and interactions between people. These are unpredictable but need to be documented
Participatory and empowering	Initiative develops in response to stakeholder participation. Evaluation is political
Cross-sector engagement	Sectors may bring different values and goals to the evaluation
Holistic and positive view of health	Broad range of positive health indicators needed to assess outcomes
Focus on equity	Impacts on equity of access and outcomes should be assessed as part of the evaluation

Empowerment of communities and building partnerships between health professionals and community members are key foci for health promotion interventions and so are important to reflect in the evaluation methods (Baum, 1998). Poland (1996a) advocates for the use of participatory action research and a critical social science perspective in deriving a conceptual model to organise evaluation and learning for healthy community initiatives. Participatory research is described as the base for evaluation in the community setting and is defined as a collective project of

researchers and people affected to produce knowledge. It is educative, draws on practice and feeds back to practice. Further, Poland (1996b) argues that evaluation is political and value-based so that elements of critical social science theory such as power, assumptions, contingency, social structures and individual agency, are all relevant to the evaluation. Thus, an important role for the evaluation is to check not only what is done but how it is done (Poland, 1996b).

An holistic approach suggests the need for evaluation at an ecologically whole system level rather than assessment of each discrete program or project on its own (Dooris, 2005). South and Woodall (2012) note a tension between an ecological, whole system approach to evaluation and the need to develop a sound evidence base. Further, positive health and wellbeing are difficult concepts to measure, in contrast to mortality and morbidity data which are better defined and comparatively easy to obtain (Hancock, 1993). Hancock suggests that a 'salutogenic epidemiology' (p17) is needed that can identify the causes of good health and the relative impact of different health determinants.

The report from the Measurement and Evidence Knowledge Network of the WHO Social Determinants of Health Commission (Kelly, Morgan, Bonnefoy, Butt, & Bergman, 2007) describes the processes, principles and challenges of evaluating action on the social determinants of health in order to address inequities. Although this report is concerned primarily with evaluation of policy rather than other interventions, it is relevant to health promotion initiatives that include strategies to address the social determinants of health and equity. Principles for developing an evidence base for policy and action are described with a focus on a commitment to equity. The report provides support for evaluation design that includes assessment of

impact on equity by, for example, health equity auditing.

So far, the literature review has charted the contested nature of health promotion, described the characteristics of community-based health promotion and presented the implications for evaluation. The next section of the literature review turns its focus to the development of evaluation of social programs generally and to evaluation of community-based health promotion programs in particular.

2.4 Developmental history of evaluation

This section presents a critical history of the development of evaluation and its relevance to community-based health promotion initiatives. The section is arranged chronologically from the beginnings of modern evaluation, with a focus on quantitative measurement and methods. Since then evaluation has undergone a number of developments as different theories and methodological approaches have evolved. Evaluation practice now reflects these developments and ranges from a technical measurement exercise through descriptive, judgemental, responsiveness and, most recently, theory-driven approaches. The emergence of a constructivist paradigm shifted the focus in the evaluation literature to the notion of evaluation as a collaborative enterprise aimed at uncovering multiple understandings of an intervention to be shared by all stakeholders. Early attempts at laying out the inter-related steps between inputs, outputs and outcome revealed the lack of underpinning theory for most social interventions and this led to the development of theory-driven, or theory-based, evaluation (Birkmayer & Weiss, 2000; Judge & Bauld, 2001). Many interpretations of program theory, and program logic models to make the theory explicit, have since been promulgated.

An approximate time line for innovation in evaluation theory and practice is shown

in Table 2.3. The timeline indicates a landmark era when the evaluation type entered the mainstream evaluation literature rather than the period of development leading to this point. The table is not intended to suggest that newer approaches have replaced older approaches; rather they co-exist within the evaluation community.

Table 2.3 Timeline of development of evaluation approaches

Landmark date	**Evaluation type**	**Theoretical perspective and focus of evaluation**
1900	Measurement	Positivist; technical measures
1930	Descriptive	Positivist; extent of goal attainment, technical measures
1967	Judgement	Positivist; development of program goals, extent of goal attainment, technical measures
1989	Fourth generation	Constructivist; dialectic and responsive
1980s	Theory-driven	Shift of focus from methods to theory, pluralist methods, sequential chain of events
1990s	Proceed-Precede Program logic models	Focus on links between planning and evaluation, sequential
1997	Realistic	Realism; what works for whom and in what circumstances
2000	Developmental	Accepts turbulence and adapts to realities of complex, non-linear dynamics

In the remainder of this section the developmental history for some major advances in evaluation thinking is outlined, with a brief description of implementation according to the main proponents. Critiques of the approach are reviewed and the application of the approach to health promotion evaluation is discussed.

2.4.1 The early years: the positivist paradigm

The positivist paradigm holds that science provides unambiguous and accurate knowledge of the world with scientific knowledge cast as objective and value-neutral (Crotty, 1998). Early evaluation was firmly based in this theoretical perspective and it remains dominant in much evaluation theory and practice.

Modern program evaluation began in the 1960s (Chen, 1990) and was grounded in positivist thinking. Weiss (1998) marks the 'War on Poverty' in the mid 1960s as the start of large scale government funded evaluation in the United States. Before then she cites some earlier attempts at evaluation in medicine and social programs, dating from 1912. By the end of the 1970s, program evaluation had become commonplace across government agencies in the United States with numerous research centres established to undertake this work. Since then the evaluation industry has waxed and waned with the availability of funding and appetite for innovation, new programs being the most likely candidates for evaluation investment (Weiss, 1998).

In Australia, early attempts at program evaluation began in the field of community education in the 1950s and 1960s and evaluation practices were adopted in social work and health disciplines in the 1970s (Sharp, 2003). In 1979 the government commissioned the Baume Report, *Through a Glass Darkly*, (Senate Standing Committee on Social Welfare, 1979) which reviewed evaluation in health and social welfare services. This report noted an almost complete absence of formal evaluation in Australian health and welfare services prior to 1973 but was able to list some 43 evaluation reports by mid-1978, mostly by government departments or commissions. The report recommended a definition of evaluation as '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, 1979 p5) and

stated its purpose was to 'provide evidence of the outcome of programs so planners can make wise decisions about those programs in the future' p6. Thus, the focus was on accountability and decision making rather than program improvement or theory building. Other sectors and jurisdictions started to develop performance evaluation but it was the late 1980s before evaluation was endorsed as a mandatory requirement of public sector programs (Sharp, 2003).

Critique of positivist evaluation and implications for health promotion

The 1970s and 80s saw the start of a shift away from evaluation firmly grounded in positivism and a move by some evaluators to more values-based approaches. In Australia, Furler (1979) was questioning the dominant approach to evaluation at that time. She described this as 'program outcome evaluation' using methods derived from experimentation and the scientific method and with a focus on the extent to which pre-determined goals have been reached. Furler (1979) argues that use of the scientific method in social program evaluation is a risk since it produces little evidence of outcomes and therefore condemns programs which may be achieving changes that are not apparent using this approach to assessment.

For Guba and Lincoln (1989) problems with a positivist approach to evaluation include that it ignores differing power relationships, the political nature of evaluation. Furler (1979) argues that the positivist approach cannot accommodate social programs since these embody ideals, a theory of intervention and implementation of the theory. All these require the setting of value criteria and making value judgements. Further, goals are often vague, ambiguous, conflicting and implicit, requiring a 'value-critical' approach to bring coherence to the program (Furler, 1979). Interestingly, while the Baume report came down heavily in favour of goal-driven evaluation, it did recognise that social program evaluations could not be

conducted in the same way as laboratory experiments and that 'withholding a program in order to provide a control group is inconsistent with natural justice' (Senate Standing Committee on Social Welfare, 1979 p97).

Another critique of the positivist paradigm is that it neglects to take account of context (Chen, 1990; Guba & Lincoln, 1989). In a community setting, variables to be measured cannot be isolated from others and the variables are likely to change during implementation (Furler, 1979). Evaluations under this positivist paradigm tend to have a narrow scope, focussing on areas for which quantitative methods work best (Chen, 1990).

2.4.2 Fourth Generation evaluation

One of the first developments to move evaluation away from a purely technical exercise was the so-called fourth generation evaluation (or naturalistic, as it was first named) (Guba & Lincoln, 1989). This approach is founded on a rejection of the positivist, scientific method and a move to constructivism.

Guba and Lincoln (1989), introducing the concept of fourth generation evaluation, describe the changing focus of evaluation and roles of evaluators over the previous three 'generations': measurement; description and judgement. It should be noted that these three approaches to evaluation are not mutually exclusive and continue to be used by many evaluators today. However, for many evaluation contexts, these earlier approaches present problems and leave gaps in our understanding.

The essence of fourth generation evaluation is that it employs a constructivist methodology and is responsive to stakeholders. Guba and Lincoln's (1989) constructivist paradigm is based on an ontology that sees reality as a social construction. Methodologically, there is a rejection of experimental/control for a

dialectic that uses interaction between stakeholders (including the evaluator) to create constructed reality. This view allows multiple perspectives from different stakeholder groups to be considered and gives recognition to the importance of the context in evaluation of a community-based initiative.

Responsiveness is achieved though interactive negotiation with stakeholder groups, to identify their claims, concerns and issues (Guba & Lincoln, 1989). These are then introduced to other stakeholder groups and resolved if possible. What is not resolved becomes the driver for data collection, using quantitative or qualitative methods as appropriate. The evaluator's role becomes that of negotiator and guide in an effort to reach consensus based on the new information gathered. Ideally, what is not resolved becomes the subject of further evaluation.

Critique of fourth generation evaluation

The major value of this approach is in the implied empowerment of stakeholder groups, not just program managers and funders, and the search for common issues to be resolved through the process of the evaluation. There is recognition of the political nature of evaluation and the different, and often competing, interests of the many stakeholders leading to different constructions about the program.

There are, however, unresolved problems with fourth generation evaluation and it remains somewhat of an ideal rather than realistic practice. Indeed, Guba and Lincoln's (1989) position with regard to constructivism has been described as 'extreme' (Laughlin & Broadbent, 1996 p435). Fishman (1992), while basically supporting the constructivist paradigm, provides a critique of fourth generation evaluation by arguing that under constructivism, it is itself a construction and therefore only one of multiple and alternate constructions. This means a special

argument for fourth generation evaluation cannot logically be made and therefore, as Guba and Lincoln themselves acknowledge, is subject to re-construction whenever new information arises. Also questioned is the assumption that action will follow from consensus in an unproblematic way, with little guidance from Guba and Lincoln (1989) about who is responsible for action (Laughlin & Broadbent, 1996).

Other critiques of fourth generation evaluation centre on the related issues of power, the role of the evaluator in a political context, stakeholder participation and practical constraints. Fourth generation evaluation requires the program manager and the evaluator to give up power to a collaborative approach where all stakeholders (including the evaluator) share their views. Abma (2005), discussing health promotion evaluation, notes that the requirement for the evaluator to relinquish control and tolerate ambiguity is presented as unproblematic but actually needs particular skills in interpersonal communication and negotiation. In fourth generation evaluation, one role for the evaluator is to gather knowledge of the social, political and cultural context through local informants but Gregory (2000) maintains that this privileges those who fill this role and can affect the evaluation process. Information flow and control is frequently subject to gatekeepers and Guba and Lincoln (1989) propose special consideration be given to ensure cooperation. However, this risks legitimising disruptive practices and perpetuating power imbalances (Gregory, 2000).

The process of fourth generation evaluation requires groups who are frequently in political conflict and with different interests to come to consensus (Fishman, 1992). This suggests a naïve view of the way values may be changed to accommodate others and underplays the fundamental conflicts (Gregory, 2000). Laughlin and

Broadbent (1996) question the ability of the negotiation process to bring about change in participants' constructions and consensus on the claims, concerns and issues since many of these are based on value differences rather than simply the need for new information.

Pawson and Tilley (1997) are particularly scathing of fourth generation evaluation and the role of evaluator as negotiator. They maintain that this fails to appreciate the asymmetry of power between different stakeholder groups. Further Pawson and Tilley (1997 p20) describe a 'deep-seated air of unreality' about the notion of evaluation as negotiation. They suggest, for example, that a joint construction of the claims and concerns of neighbourhood watch participants and local burglars is hard to imagine! Fishman (1992) agrees that the approach appears to ignore the differential power relationships, for example, program funders and managers who bring their own goals to the program. The evaluator is privileged by holding resources, status and a leader or facilitator role (Gregory, 2000) and by controlling what information is to be obtained and its interpretation (Laughlin & Broadbent, 1996). Indeed as the only person with access to sufficient data to form a well-rounded view of the situation, the evaluator is assured a position of superiority (Gregory, 2000). As Potvin and Bisset (2009) note the influence and power of the evaluator is underestimated in fourth generation evaluation.

The issue of social heterogenesis is raised by Gregory (2000) in her critique of participation in fourth generation evaluation and Fishman (1992) highlights a further issue concerning the selection of participants from potentially hundreds of stakeholders by asking who to engage with and who has priority? While Guba and Lincoln (1989) attempt to deal with this by suggesting that the evaluators should

make every effort to engage with and take into account all stakeholder groups, they also recognise that this process of inclusion may be restricted by practical constraints such as a project's resources.

Also in terms of practical constraints there is a lack of guidance on implementation of fourth generation evaluation. Fishman (1992) notes there are no details on implementation or case studies to guide the evaluator and the process of implementation as described by Guba and Lincoln (1989) is unwieldy and difficult. In Abma's (2005) example of fourth generation evaluation of an injury prevention program in a dance school, the junior (student) evaluator spent three to four days per week for a whole year in attending lessons, consultations and performances in order to build relationships with the school community. Case study examples conducted in a more open system such as a city suggest that six months is needed to gain insight into the setting and meet with stakeholders before the focus of the evaluation is identified (Abma, 2005).

Finally, Gregory (2000) points to a conflict between Guba and Lincoln's (1989) view that evaluation design should be a continuous, emergent process, and their recommendation for a contract that estimates events, resources and products to satisfy funders and managers. This puts constraints and boundaries on the evaluation that appear to be in conflict with the emergent design of fourth generation evaluation.

Implications for community-based health promotion evaluation

Despite the issues critiqued above, the fourth generation evaluation approach has much to offer community-based health promotion in terms of freeing the evaluation from the tyranny of a positivist paradigm that requires controlled experiments to demonstrate causal outcomes. For health promotion interventions that are designed to

be participatory and empowering of individuals and communities, this evaluation approach at least attempts to place participants as equal partners although, as noted above, there are unanswered issues of power and control. The recognition of different interests, values and understandings (constructions) of the health promotion program reflects the political nature of health promotion and the involvement of multiple disciplines and sectors. As Abma (2005) notes fourth generation evaluation is synergistic with health promotion in that it recognises active participants rather than passive research objects, it can accommodate multiple interacting factors and perspectives and moves from professional dominance to shared decision making. The evaluation is designed as a continuous and emergent process which is appropriate for a developmental program that needs to be flexible to local needs and interests. However, as described in the critique above, issues of power differences, particularly between program funders, managers and practitioners and the lay community participants of a program, remain largely unchallenged, as does the assumption that consensus can be obtained and action will follow. Also problematic is the implicit timeframe and resources required to conduct a fourth generation evaluation. While this might be feasible with a well-funded, long-term national program, it is unlikely to be realistic for evaluation of local initiatives that do not have funding for a comprehensive, knowledge development style of evaluation but are required to report to funders and policy makers about the outcomes of their program.

Some of these issues, in particular the political nature of evaluation and the problem of generalisability of fourth generation research findings is an ongoing concern and one that the next development – theory-based evaluation – attempts to address.

2.4.3 Theory-based evaluation

Theory-based evaluation is a relatively recent development. Previously to this,

evaluation was most often a-theoretical and driven by methods (Chen, 1990). The focus was on inputs and outcomes with a 'black box' of unknown mechanisms and processes with no analysis of how these might be linked (Scriven, 1981). Program theory and program logic modelling in various guises have now been widely adopted in social science research and evaluation (Birkmayer & Weiss, 2000; Rogers, Petrosino, Huebner, & Hacsi, 2000; Rossi, Lipsey, & Freeman, 2004).

Ecological evaluation, outcomes hierarchies, and use of program theory and program logic models are three inter-related approaches to theory-based evaluation and are discussed below.

Ecological evaluation

Ecological models are characterised by attention to behaviour and to individual and environmental determinants that show reciprocal causation (McLeroy, Bibeau, Steckler, & Glanz, 1988). Ecological evaluation is based on an understanding of that reciprocity, that is, behaviour modifies the environment and that the environment modifies behaviour (Green, et al., 2000). This approach also recognises different levels of analysis for environmental influence such as individual characteristics, personal networks of family and friends, institutions and the relationships between them, and public policy/laws (Dooris, 2005, 2012; McLeroy, et al., 1988).

The Precede-Proceed ecological evaluation model (Green and Kreuter 1999) is an early attempt at making clear the links between health education program planning and evaluation and the series of steps each of these requires. It introduces the ideas of progressing outcomes from short-term to intermediate-term to long-term (see Figure 2.3). Green and Kreuter suggest that the Precede-Proceed model produces a series of hypotheses about presumed relationships between interventions and systems.

Evaluation then provides a 'test' of the hypotheses.

In the Precede-Proceed model, Phases 1 to 5 represent the planning and assessment stages, leading to Phase 6 Implementation. Phases 7 to 9 form the evaluation component.

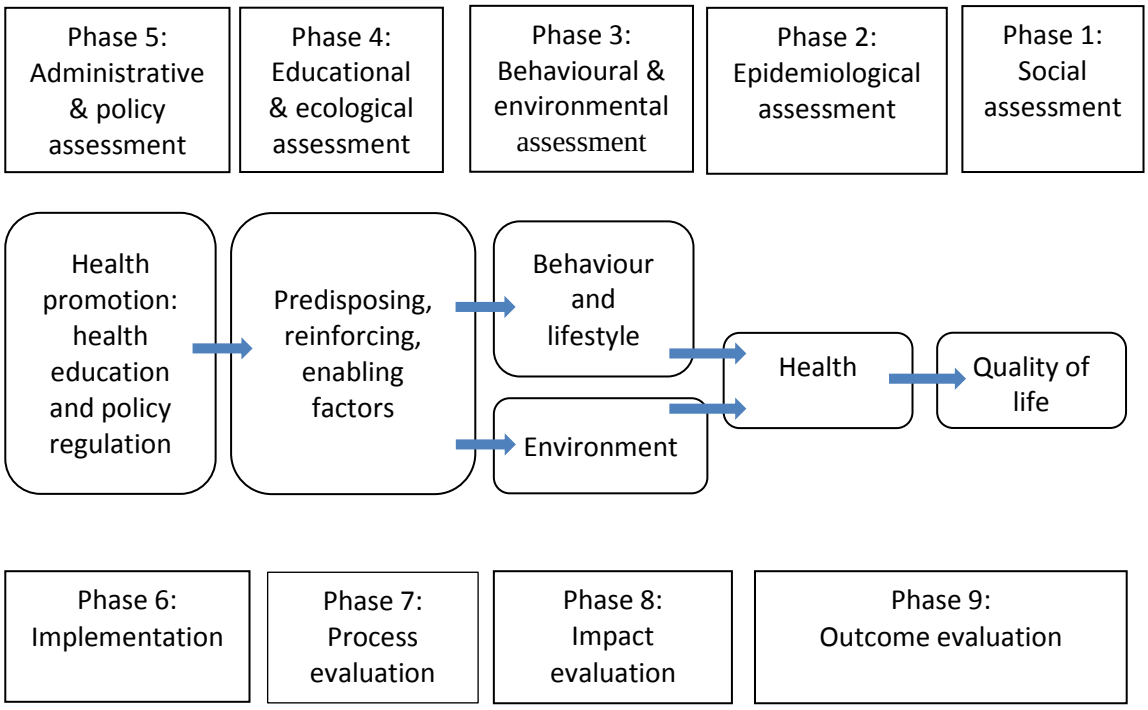

Figure 2.3 Stages of planning and evaluation in the Precede-Proceed model (based on Green and Kreuter 1999 p35)

The Precede-Proceed model has been influential in health promotion. For example, Jones and Donovan (2004) report on a survey of Australian Health Promotion Association practitioner members that found 78% of respondents claimed familiarity with Precede-Proceed (and this was the most familiar health promotion theory/model) and just over half of these used it in their practice. Of the 30% that claimed to use any theory or model in evaluation, Precede-Proceed was most often

mentioned. Despite this, there are few documented examples of use of Precede-Proceed in community-based health promotion evaluation. Although the website (Green, undated) reporting uses of Precede-Proceed cites many hundreds of studies, a review of the report titles indicates that these mainly focus on the planning component of the model. The small sample of reported uses in evaluation are mostly clinical or health education interventions.

One reported community-based example of Precede-Proceed is in the evaluation of a child pedestrian injury prevention program (Howat, Jones, Hall, Cross, & Stevenson, 1997). The authors claim that the benefit of using Precede-Proceed is that 'the likelihood of a rigorous evaluation design is enhanced' (p 286). A flow-chart linking objectives and strategies with the social, epidemiological, behavioural and environmental factors was identified and this was used to monitor program implementation. Howat and colleagues (1997) conclude that using Precede-Proceed was 'onerous' but enhanced the quality of planning. It is not clear how the Precede-Proceed-based flow-chart contributed to evaluation beyond evaluating the process of implementation. Another reported evaluation of a youth mental health awareness campaign used the Precede-Proceed model to guide development, implementation and evaluation (Wright, McGorry, Harris, Form, & Pennell, 2006). The authors conclude that the use of Precede-Proceed contributed to the effectiveness of the strategy by fine tuning the campaign targets through the population assessment stage.

The main strength of the Precede-Proceed model is the clear link provided between the planning and evaluation stages and the description of the steps between planning, implementation and outcomes of an intervention. This encourages program planners to assess the various factors that have potential impact on the program. The joint

influence of behaviour and environment and the different analytical levels described in the phases support the notion of an ecological model. Limitations are the linear nature of the model and the lack of feedback mechanisms. There is no explicit theory articulated and Green and Kreuter's main focus is on health education programs rather than broader definitions of health promotion. Another limitation is that it does not address questions of generalisability or causation. Green and Kreuter (1999) ask about 'internal validity' – does the research design support claims that results really stem from the program? and 'external validity' – how generalisable is the program? But the model does not explain how these questions might be answered. It is important to note that 'what 'proceeds' is not the same as 'what causes'. As with earlier evaluation methods, establishing generalisability of findings and causation remain elusive.

In Australia, Hawe and colleagues (1990) expanded the concept of the Precede model beyond health education to health promotion more broadly. Their popular book is aimed at health promotion practitioners rather than evaluators, reflecting the authors' belief that process evaluation and evaluability assessment are best conducted by program staff. The expertise of evaluation research staff is called upon at the point of assessing short and longer term program effects. Like the Precede model, Hawe and colleagues (1990) set out a pathway for program planning and evaluation starting from a measurable goal, measurable objectives and sub-objectives, and strategy objectives and activities. Process, impact and outcome evaluation measure the activities, the immediate impact and the long-term effect respectively. The notion of a planning and evaluation cycle (see Figure 2.4) is introduced, stressing the need for continuous monitoring during implementation and also introducing the concept of feedback loops from outcome evaluation back into

needs assessment.

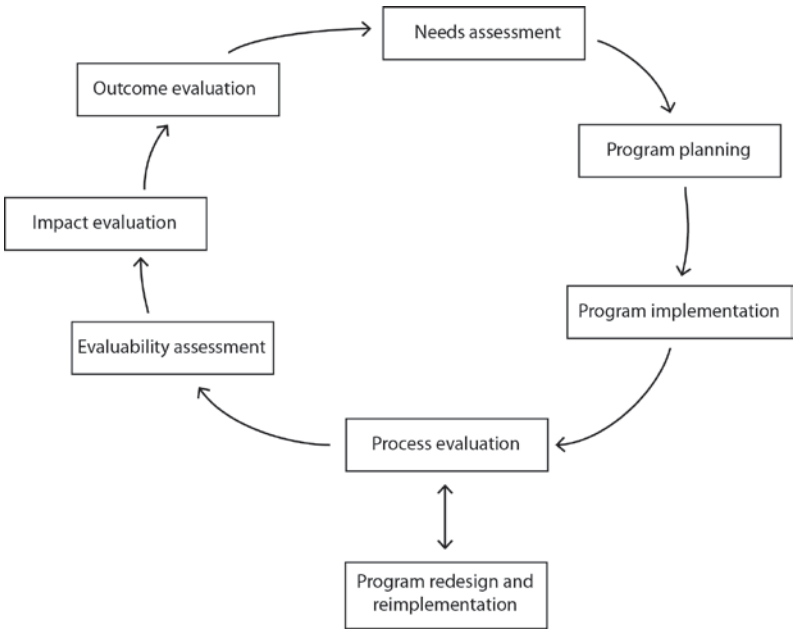

Figure 2.4 Planning and evaluation cycle (based on Hawe, et al., 1990 p78).

Evaluability assessment is described as providing a framework for decisions about what aspects to focus the evaluation on, who needs the evaluation information, how the evaluation will be conducted and what measures will be used. The evaluability assessment process includes clearly defining the program (getting consensus on program boundaries), specifying goals and expected effects (along with unintended effects) and ensuring the plausibility of causal assumptions in the program. The pre-condition for evaluability assessment is a 'rational fit between clearly defined programme activities and the programme goals' (Hawe, et al., 1990 p89). This suggests a search for program theory as a precursor to the evaluation effort.

Hierarchy of outcomes

The second approach considered under theory-based evaluation is the use of a hierarchy of outcomes to articulate the links between activities and outcomes at different timeframes. Outcomes from health promotion can span across timeframes

from short to long. Drawing on the notion of steps or staged outcomes, a hierarchy of health promotion outcomes illustrates levels of achievement according to the timeframe for the program. Each stage leads on to, and is a pre-requisite for, the next. WHO (1998) definitions of immediate, intermediate and health outcomes emphasise the attributable outcomes of planned interventions and that these changes may be in individuals, groups or whole populations. In using the outcomes hierarchy model, evaluators need to consider at what level (that is proximal/ immediate outcomes or more distal long-term outcomes) the evaluation can be conducted. This will depend on factors such as available resources, skills and timeframe.

Nutbeam has developed an outcomes hierarchy that explores the different stage of a health promotion program (Nutbeam, 1996, 1998). This approach helps to identify what might be realistic for a health promotion program to achieve given its time frame and resources, and may resolve some of the challenges in evaluating short-term funded programs for long-term results. Immediate, health promotion outcomes are changes in individual attitude, knowledge, skills and behaviour or changes in organisational practice or public policy that will lead to a more health promoting environment. (Nutbeam, 1996, 1998). They are generally assumed to take up to two years to achieve. These changes are relatively easy to measure in individuals using psychometric scales, or, in the case of skills and behaviour, by observation. Organisational and policy change can be assessed through documentation or survey of stakeholders. To strengthen the evidence that change has taken place, baseline data should be available or collected in the planning stage of a program.

The next stage of intermediate outcomes represents changes to the determinants of health. Examples include changes to lifestyle, increased access to effective health

services, increased food security and more planned health promoting urban environments (Nutbeam, 1996, 1998). These outcomes might take two to five years to achieve. Intermediate outcomes are more difficult to assess directly since they are conceptually complex and have multiple interpretations. Indicators may be developed that operationalise the outcomes into measurable changes. Examples of indicators are: use of health services by marginalised groups compared to the general population; cost of a standard basket of food compared to average wage; hectares of green open space per unit of population. It is here that attribution issues become more problematic.

The highest outcome level describes changes in health status or in more equitable health status; that is improved quality of life, or reductions in mortality and morbidity (Nutbeam, 1996, 1998). Typically these outcomes could take ten years or more. Examples of long-term outcomes include increased life expectancy, reduction in lung cancer rates and decreased child mortality. A long-term health outcome of particular relevance in most developed nations is for an increase in the proportion of people of healthy weight and a decrease in prevalence of overweight and obesity.

While the outcomes described above may be measured directly or indirectly, it is clear that the further along the distal chain the more difficult it becomes to attribute an observed change to an intervention. Program theory, considered next, attempts to provide plausible and logical links between the outcome levels in order to provide support for causal links.

Program theory and program logic models

The notion of program theory represents a critical step forward in describing and understanding how and why a program works or does not work as was expected. It

encourages the evaluator and other stakeholders to examine more closely the theoretical assumptions linking the program components and underpinning the relationship between activities and expected outcomes. Program theories arise from social science theory or the logic of a program logic model (Birkmayer & Weiss, 2000).

An oft-cited definition of program theory is ‘a plausible and sensible model of how a program is supposed to work’ (Bickman, 1987 p5). Chen (1990 p43) broadens this definition to ‘a specification of what must be done to achieve the desirable goals, what other important impacts may also be anticipated, and how these goals and impact would be generated’. Funnell and Rogers’ (2011) definition ‘theory or model of how an intervention contributes to a chain of intermediate results and finally to intended or observed outcome’ (p xix) includes the notion of causation by noting that the program theory describes how the intervention contributes to outcomes. Weiss (1998) has a simple definition of theory as ‘a set of beliefs that underlie action’ p55. She reminds us that a theory does not have to be universally accepted, or even right, but consists of a set of hypotheses upon which programs are built. The purpose of the program theory is to clarify and develop the stakeholders’ theory or program (Chen, 1990) and as a tool to help decision-makers and others to develop policy and evaluation questions (Grembowski, 2001).

While program theory and program logic model are terms that are often used interchangeably in the evaluation literature, they do have distinct meanings. These are summarised in Table 2.4.

Table 2.4 Program logic models and program theory

	Program logic model	**Program theory** **In addition:**
Description	Graphical representation of program components i.e. inputs, activities, outputs, outcomes (Chen, 1990)	Model of how the program is expected to work (Bickman, 1987) Systematic configuration of prescriptive and descriptive assumptions underlying a program (Chen, 1990)
Purpose	Evaluability assessment (Chen, 1990) Clarify and make explicit the links between program components (Mark, 2003)	Identifies theoretical assumptions underpinning program and the mechanisms expected to have an effect (Mark, 2003)
Evaluation role	Monitor performance against expectations (Chen, 1990)	Soundness of program and how it works (Chen, 1990) Explains how mechanisms are causally linked to outcomes (Funnell & Rogers, 2011)

Mark (2003) distinguishes between a program logic model (inputs, activities, outputs, short and long-term outcomes) and program theory which, in addition, identifies the mediators through which the program is expected to have effects, that is the mechanisms triggered to produce the outcomes. Thus program theory, while it may include a logic model depicting the causal chain, takes this a step further by identifying the underlying theoretical assumptions about the causal chain. Program

theory can be used to show where chains of assumptions are well-supported by the data and which chains break down. This should lead to the development of different and, hopefully, more effective strategies as these gaps in the theory are identified (Birkmayer and Weiss 2000). Program theory can help to distinguish program failure from implementation failure and identify the connection between activity and effect (Chen, 1990). Thus, the program theory should be able to explain the program-activated mechanism which has a causal link to the outcomes.

This potential to explain causal mechanisms of change may assist in generalising to a broader range of settings. Program theory facilitates generalisation, because if it is known how and why a program works, (that is the mechanisms underlying the program effect) this helps to assess if it will do so in new situations (Mark, 2003). An example of this would be the expansion of the Healthy Cities approach to Healthy Islands, Healthy Markets and so on. As Birkmayer and Weiss (2000) note each evaluation is a 'prisoner' of its context but may build the body of knowledge about what actions lead to the desired change. The evaluation can then prioritise those parts of the chain where uncertainty about causal links is highest and where reducing uncertainty could make the most difference (Patton, 1997).

Program theory can be divided into two parts; 'theory of action' and 'theory of change'. Theory of action describes how the program works (Funnell & Rogers, 2011) and what it does (Chen, 2005). One of the first models to describe a program's chain of events as a theory of action is Bennett's 1979 model – including a hierarchy of inputs, activities, participants, reaction, change in knowledge, attitudes and skills, change in behaviour, longer term outcomes (Patton, 1997). Explanation of how the change comes about is described by the 'theory of change' (Chen, 2005; Funnell &

Rogers, 2011) and can also explain the plausible, logical links between program components (Judge & Bauld, 2001; Mackenzie & Blamey, 2005). The conceptual framework for a theory-driven evaluation is illustrated in Figure 2.5.

Existing or implicit theory leads into the theory of action which describes the planned workings of the intervention. Following implementation, the theory of change describes how the intervention is believed to have worked (assuming outcomes are observed). Environmental influences can have an effect at the theory of action or theory of change stage.

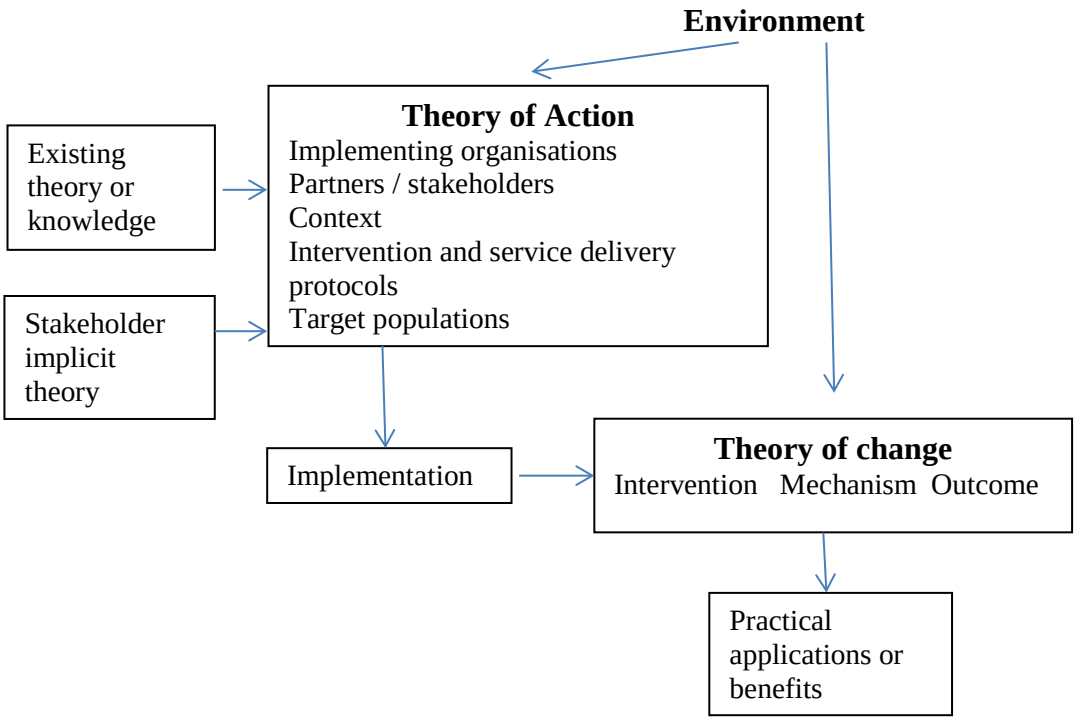

Figure 2.5 Conceptual framework for theory-driven evaluation (adapted from Chen, 2005).

Theory-based evaluation is said to bring to the surface the assumptions upon which the program is based, that is the sequence of steps from activity to outcomes, or a series of 'little theories' (Birkmayer & Weiss, 2000). So evaluation data is needed not just on the achievement of outcomes but also to question the steps along the way

and to ask ‘are the theories realised in action?’ (Birkmayer & Weiss, 2000). In this way theory testing becomes a way of unpacking the evaluation ‘black box’ and should explain how and why programs achieve or don’t achieve (Weiss, 1998) (Birkmayer & Weiss, 2000), in contrast to just ‘do they work?’

Patton (1997) proposes a ‘user-focused approach’ where the evaluator’s task is to facilitate articulation of the operating theory (Chen’s descriptive theory) or theory of action by intended users. So for Patton, the program theory is generated by the stakeholders rather than through deductive or inductive methods. Further, Argyris (1982) distinguishes between espoused theory (what people believe and say) and theory-in-use (how people act). In a user-focused theory of action approach the evaluator needs to assist with drawing out espoused theories and testing these against actual theories, leading to reality testing – asking whether what the stakeholders believe to be the case is what actually occurs (Patton, 1997). In health promotion, for example, provision of health education in order to change behaviour may be based on a theory that more information leads to a healthier lifestyle. In reality, a host of other competing theories could contribute to explaining what changes behaviour. This underlines the complementary nature of the logic model and the program theory, as articulation of the links in the logic model encourages stakeholders to consider what theories and assumptions are being used to develop the underlying program theory.

A number of theorists (see, for example, Birkmayer & Weiss, 2000; Chen, 1990; Donaldson, 2003; Grembowski, 2001) have described the steps to be undertaken in conducting a theory-driven evaluation. Program theory needs to be articulated early in the evaluation process (Birkmayer & Weiss, 2000) and this planning stage

develops the program theory from multiple sources of information that may include: prior theory; implicit theories; observations of the program in action; or exploratory research to test critical assumptions about the nature of the program as the program theory is developed (Donaldson 2003). The next step is to use the program theories to develop evaluation questions in consultation with stakeholders (Donaldson 2003). The theory-driven approach is described as 'method neutral' and methodological choices are informed by program theory, by the evaluation questions and by practical constraints (Chen, 2005; Donaldson, 2003).

Critique of theory-based evaluation

Program theory aims to make clear the underlying assumptions about how the program is expected to work but it may reveal that different stakeholders hold different beliefs about this (Grembowski, 2001; Weiss, 1998). However, it is important to remember that theory provides a guide to the evaluation and is not always correct (Birkmayer & Weiss, 2000) or is only partially correct, poorly developed or there are competing theories about the same program (Birkmayer & Weiss, 2000; Mark, 2003). For example, Birkmayer and Weiss, reviewing the evaluation of a pregnancy prevention program, suggest a number of possible theories: the program provides information about contraception leading to increased use; it teaches young women to be more assertive about demands on their partners; it makes chastity more socially acceptable within the program group. In contrast to this multiple theory approach, program logic models tend to be simplistic and linear (Poland, 1996a; Rootman, Goodstadt, Potvin, & Springett, 1997). Further, they represent the program at one point in time and can miss unintended consequences. For this reason Poland (1996a) suggests models need to be re-visited as the program unfolds.

Another issue for critique is the notion of the participatory process in theory-based evaluation. While there seems to be nothing fundamental about this approach that dictates a participatory and consultative model, this is widely considered as a necessary part of the process. This is consistent with the earlier described fourth generation approach and with health promotion principles that stress participation and empowerment. However, theory-driven evaluation steps back from the purely constructivist paradigm of fourth generation evaluation. For example, Chen (1990) argues that, to increase objectivity, the evaluator should also use their own expertise and knowledge to develop the evaluation questions because stakeholders may miss causal processes and have vested interests. Moreover, Grembowski (2001) suggests that the evaluator should cross check that there is a good fit between program theory, objectives and evaluation questions before proceeding to the next stage. Thus, this approach acknowledges the position of the evaluator as an 'expert' and privileges this role over other stakeholders, if only to ensure that the evaluation moves forward.

The process of reaching consensus among a group of stakeholders with their own, possibly competing, interests is likely to be contentious. Normative evaluation is one approach that helps stakeholders identify and clarify program outcomes and goals (intended and unintended) (Chen, 1990). While this may be a practical approach to assist in achieving consensus, it would appear to put boundaries around the evaluation and potentially exclude some views. Also for practical reasons, the composition of the stakeholder group is likely to be limited so not all stakeholders can participate. Patton (1997) asserts that the evaluator should aim for fairness and pluralism in selecting representatives, however this approach inevitably means that the views of a range of parties are overlooked and expertise may be lost (Gregory, 2000). Further, the capacity and motivation of stakeholders to actively participate in

theory explication and testing seems to be assumed in the theory-driven approach. Poland (1996b) suggests that power-brokers may be in a position to impose their own views and Gregory (2000) points to the risk that those with more powerful voices will be disproportionally heard compared to other community members unless the evaluator is highly skilled in facilitation and takes an ethical stance to engage those with less power. Again this elevates the power of the evaluator, and relies on those with least power being willing, and enabled, to participate.

Implications for community-based health promotion evaluation

For community-based health promotion evaluation, the theory-driven approach provides an opportunity for stakeholders with different understandings of a program to participate in reaching a consensus position and articulating shared goals. The process of building the program theory provides an opportunity to strengthen a program by identifying and considering the underlying assumptions, potential areas of breakdown and unintended outcomes. The development of a program theory is likely to be useful since many health promotion programs are funded, planned and implemented without adequate attention to underlying theory (Birkmayer & Weiss, 2000; Judge & Bauld, 2001; Nutbeam, et al., 2010). However, power differentials between stakeholders remain and most commentators address this by giving the final say to the evaluator as the arbiter of program theory. Of course reaching a consensus position will be easier if stakeholders already share an understanding of health promotion principles and values as this may provide common ground and mutual interest.

Problems with the theory-driven approach include unresolved questions about the ability of program theory to explain the program mechanisms and the causal links to the outcomes, issues of representation and participation of stakeholders, and the

resource-intensive nature of the process (for funders, evaluators and participants). Barriers of resources and time constraints are particularly pertinent to evaluation of community-based health promotion, where funding is already scarce (Baum, 2003) and theory is underdeveloped (Nutbeam, et al., 2010). Participants, or intended recipients, of the health promotion programs are often from disadvantaged backgrounds and so are likely to be those with least resources to engage actively in the evaluation.

The WHO Measurement and Evidence Knowledge Network report (Kelly, et al., 2007) strongly supports the use of theory to develop a logic model or causal pathway, stating:

> 'In summary it is important to specify three things with respect to any intervention, action or policy. First, be as specific as possible about its content in its application on the ground. Second, clarify what is to be done, to whom, in what social and economic context, and in what way. Third, articulate the underlying theories which make explicit the assumed causal links between action and outcomes'. p92

However, using a logic model in this way would seem to suggest a top down and fixed intervention with little room for participatory processes or ongoing development in response to issues arising during implementation.

Hawe and colleagues (2009) provide a critique of theory-based evaluation specifically from a health promotion and illness prevention perspective. They argue that, in its attempt to pin down causal pathways and understand implementation and outcomes, theory-based evaluation has led to improved means to unimproved ends; that is a more thoughtful way of doing the same things rather than actually thinking about evaluation in a new way. Instead, they argue for an ecological systems view that is not just multi-level but also recognises system dynamics, that is, linkages, relationships and interactions. This is similar to complexity theory, which is

discussed later in Chapter Four.

In summary, theory-driven evaluation could therefore be seen as a middle ground between the constructivist fourth generation approach (which is difficult to implement in typically small scale and time-limited health promotion evaluations) and the positivist approach that demands pre-determined hypothesis testing and neglects the influence of the context or setting on the program.

2.4.4 Realistic evaluation

The next major evaluation development to be considered is realistic evaluation, which is based on the understanding that context is a critical factor in evaluating an intervention.

Development of realistic evaluation

In 1997, Pawson and Tilley published their landmark book 'Realistic Evaluation'. This details the key components of a realistic evaluation as 'context' 'mechanism' and 'outcome' and the interactions between these. The authors emphasise the importance of context and of asking not just what works but 'what works for whom and under what circumstances'. Pawson and Tilley (1997) argue that social interventions work through the action of mechanisms and that statistical or controlled research methods miss out on the explanatory ingredients, that is, the mechanism. Mechanisms are described as theories, made up of resources and reasoning, that trigger an outcome. The basis of a realist explanation of how a program operates is that causal outcomes are triggered by a mechanism acting in a context (Figure 2.6).

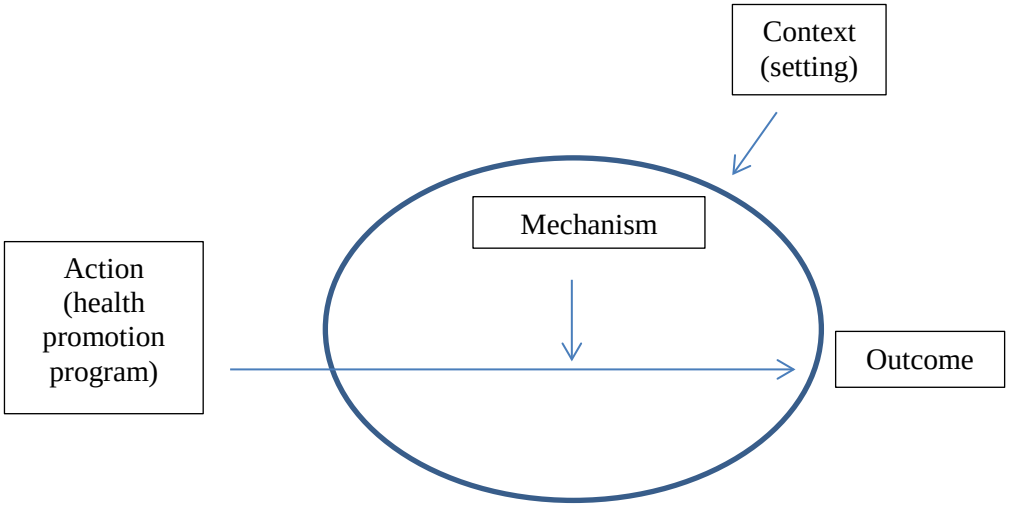

Figure 2.6 Generative causation (adapted from Pawson and Tilley 1997)

Using a generative theory of causation, rather than producing outcomes directly, the (health promotion) program is seen as an opportunity that offers participants a chance to make (constrained) choices. Thus, in this view, the program or intervention does not itself produce outcomes but generates the underlying mechanism that causes individuals or communities to change. It is the action of stakeholders that make a program work; its causal potential is to provide reasons and resources to enable participants to change. The evaluation question is then what conditions allow for this to occur and has it happened in practice?

The realist evaluation cycle is depicted in Figure 2.7 below. The starting point is the theories which are developed collaboratively with stakeholders. 'Realistic evaluation involves the researcher learning the policy, practitioner and participant ideas that constitute the program and govern its impact' (Pawson & Tilley, 1997 p207). These theories are used to elicit the context-mechanism-outcome configuration, which leads to hypotheses about how the program works. Data are collected to test, confirm or refine the hypotheses. Like theory-driven evaluation, evaluation methods are chosen

that are best able to test the hypotheses present in the context-mechanism-outcome. This leads on to program specification rather than the traditional goal of generalisability. Finally, the new context-mechanism-outcome configuration and theories feedback to stakeholders as the learning from the evaluation.

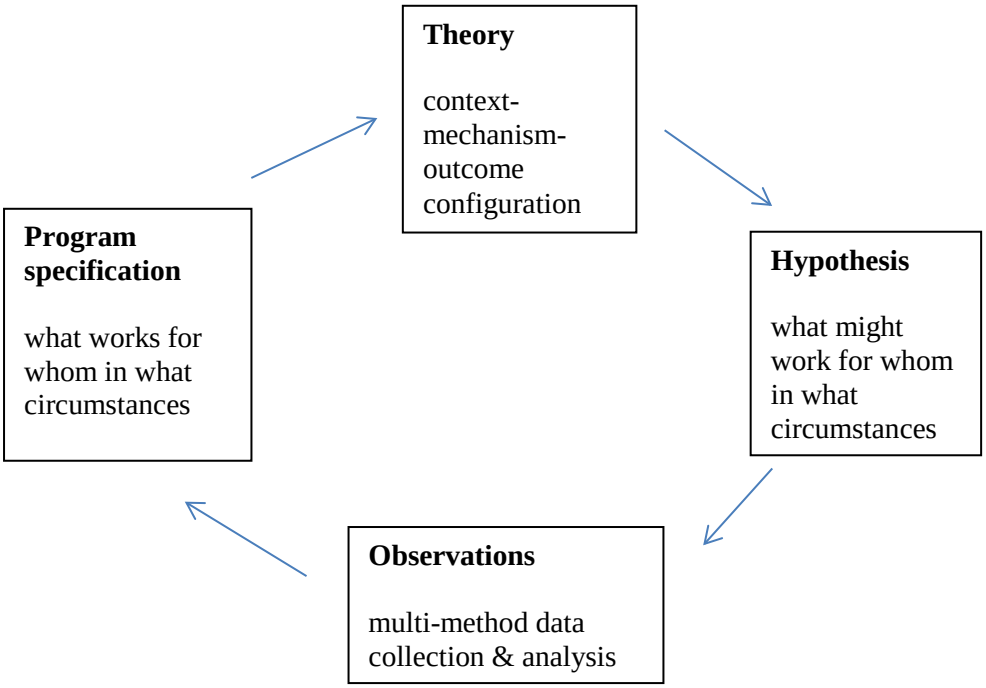

Figure 2.7 The realist evaluation cycle (adapted from Pawson and Tilley 1997)

Pawson and Tilley (1997) argue that generalisability can still be built up by developing middle-range theories from cumulative evaluations to interpret similarities and differences between programs. For example, a crime prevention project based in a housing estate in the United Kingdom produced data on burglaries showing dramatic reductions over three years. Two similar projects were conducted in other areas with mixed results and Pawson and Tilley (1997) point out that the original project could not truly be replicated due to differences in the settings. However, going back to the data from the original project revealed that re-burglary within a short space of time was a major problem and this indeed had been one element in the context of the project. International data revealed re-victimisation

following burglary was high immediately following an incident and faded by six months. Further studies and projects then followed looking at repeat victimisation across a range of crimes and these all confirmed the pattern. This led to development of a middle-range theory that 'repeated co-presence of a particular motivated offender and suitable victim in the absence of capable guardian'(context) plus an 'offender informed that, pro tem, risks low enough, rewards available sufficient and crime easy enough' (mechanism) leads to 'high rate of repeat offences with short-term heightened vulnerably' (outcome) (Pawson & Tilley, 1997 p139). This allows translation into a context-mechanism-outcome theory in which action can be targeted at the offender, the victim or the guardian e.g. measures to enable apprehension of repeat offenders, or increased security measures immediately post-burglary.

Critique of realistic evaluation

Gregory's (2000) critique of realist evaluation is based on the apparent separation of the practitioner/researcher from the subject of the intervention. Implementing realistic evaluation requires the generation of hypotheses but the method of generation is limited to the research literature and practitioner knowledge, only two of the three types of knowledge as classified by Reason (1994). Thus, the third knowledge type, experiential knowledge, is excluded and the findings are not grounded in the personal experiences of any of the stakeholders (Gregory, 2000).

Two examples of the use of realist evaluation reported in the literature indicate that identifying and agreeing on the context-mechanism-outcome configuration is more challenging than suggested by Pawson and Tilley. Greenhalgh and colleagues (2009) report on an evaluation of a health service change strategy. In practice it was difficult to identify the mechanisms for change as, although stakeholders were generally clear about long-term goals, the intermediate steps were rarely articulated.

There was a tendency for mechanisms to be become clearer in retrospect after outcomes had been identified and agreed upon. The other main conclusion from this study was that the generative causality of a context-mechanism-outcome configuration is not based on logical deduction but is an interpretive task involving much negotiation.

Byng and colleagues (2005) describe the use of realistic evaluation in assessing a set of mental health care linkage projects. The project aimed to develop systems in primary care and links to specialists in order to improve care for patients with long-term mental illness. The evaluators constructed case studies for twelve general practice sites (including control sites), from interviews with practice and program staff and analysed these to create middle range theories. Some challenges are described: for example, a multiplicity of context and mechanism groupings was identified, and in many cases it proved difficult to allocate a factor as context or mechanism. Both positive and negative feedback loops between outcomes and mechanisms complicated the process of establishing the context-mechanism-outcome configurations. Byng and colleagues (2005) conclude that in practice these difficulties were only technical in nature and the context-mechanism-outcome configuration search did help to improve understanding of program.

Implications for community-based health promotion

The importance attached to understanding the program context complements the health promoting approach and suggests a way to generalise learning across different interventions. However, community-based health promotion initiatives are likely to be working with a complex mix of underpinning theories, strategies and participating groups. While it is apparent that a program can be operating under a number of different or overlapping context-mechanism-outcome configurations at any one time,

realistic evaluation does not make clear how these different configurations can be linked or networked in a dynamic holistic context. Each scenario appears to be independent in terms of the evaluation.

The issues described above point to a deficit in current evaluation theory and practice that can do justice to community-based health promotion. The risk in continuing to use inappropriate evaluation approaches is that evidence of effectiveness from high cost health promotion trials will continue to be weak (Hawe, Shiell, & Riley, 2004; Ling, 2012) and this will contribute to the marginalisation by funders and policy makers of the health promotion sector. Some ways to address the challenges of evaluation of health promotion initiatives are outlined below.

2.5 Addressing the challenges to evaluating community-based health promotion initiatives

The main challenges to evaluating Healthy Cities and similar community-based health promotion initiatives, as identified by this literature review, comprise the following main issues: i) complexity of these initiatives and their settings, ii) using appropriate research approaches and iii) causality and attribution of outcomes.

2.5.1 Complexity

Community-based health promotion initiatives are generally set in dynamic, complex systems with each setting functioning as an open system in exchange with the wider environment and other settings (Dooris, 2005). Healthy settings approaches tend to have long-term goals, multiple actions and expected outcomes at multiple levels and are active in local contexts that differ from city to city (Baum, 2003). Further, many of the social processes underpinning action, such as empowerment and community participation, are poorly theorised or are contested in meaning (Baum, 2003). While

the use of program theory and program logic models goes some way to addressing these issues, this approach is not able to deal with the multiple layers, interactions and complexity that characterise much health promotion. Chapter Four considers the potential for complexity theory to contribute to evaluation of such community-based health promotion initiatives.

2.5.2 Research approaches

As we have seen above, research and evaluation approaches that can deal with complexity are under debate and review. The notion of the superiority of the randomised controlled trial and other experimental methods has been challenged and a mix of quantitative and qualitative methods to suit the specific evaluation question is proposed by many commentators (see, for example, Baum, 1995; Judge & Bauld, 2001; Nutbeam, 1999). As noted earlier, setting up control communities and keeping them uncontaminated by the 'intervention' is not possible or practical and since the initiative is likely to be developmental it is not usually possible to predict the intervention in advance (Baum, 2002). However, much evaluation is grounded in positivism and there is still debate about the validity of qualitative approaches. So, for example, writing from a health promotion practitioner viewpoint, Smithies and Adams (1993) claim that evaluators are usually academics immersed in the dominant scientific paradigm, using methods that are inadequate and inefficient. Thus, Smithies and Adams (1993) argue that evaluators should aim to work in partnership with the subjects of the evaluation and create new methodologies that are more appropriate for a health promoting approach. This is clearly a call to separate health promotion evaluation from the medical model and to strengthen participatory approaches.

One of the positive newer developments in evaluation is the prominence given to

environmental contexts when assessing program effects. This means that in health promotion evaluation, outcomes are not restricted to individual behaviour change but can also include changes in the social, legislative and political environment.

Also promising is the development of theory-based evaluation and its many offshoots. These approaches stress the importance of using best fit methods for gathering data to address the evaluation questions, rather than following a particular quantitative or qualitative method. Thus, health promotion evaluators need to be method pluralists while considering the values and needs of the program stakeholders, as well as the available resources and the practical constraints in conducting a community-based evaluation. Resources may constrain the extent of the evaluation so consultative choices about where to focus effort will likely need to be made.

2.5.3 Attribution and causality

These two issues of complexity and use of non-experimental methods leads to the conclusion that a linear model of causality cannot be established with any certainty. The long time frame required for achieving outcomes from initiatives such as Healthy Cities adds to the problems of causality and attribution of effect.

Green and Kreuter (1999) argue that evaluation should include analysis of actions pertinent to controlling the determinants of health or quality of life issues that have been previously identified as relevant to the problem addressed. There should also be environment assessment of factors that 'could be causally linked' to the behaviour change identified (Green & Kreuter, 1999 p112-113). Thus, previously identified research evidence can be used to strengthen claims of causality.

Theory-driven evaluation aims to make explicit how and where causal pathways are

predicted. In this approach, a theory of causality is tested and if the underlying program logic stands up this is taken as evidence that some degree of causality is present. Realistic evaluation takes this a step further by using a generative theory of causation that replaces the need for a causal link to a direct outcome and instead considers the causal power of a program to enable change in the participants. In health promotion initiatives, then, the evaluator needs to understand the program theory and the mechanisms by which the activities of the program are expected to bring about change, and the setting or context in which the program unfolds.

2.6 Chapter summary

This chapter has described the history and development of health promotion over the last four decades. Over this time, debates have continued about the effectiveness of the two broad approaches to health promotion practice. The first, based on social marketing and education theory, leads to health promotion initiatives such as mass media campaigns to raise awareness of a health issue and educational materials aimed at increasing knowledge and skills in this area. These activities are expected to produce behaviour change that increases healthy lifestyles and decreases individual risk factors, and hence improves health. The second approach uses socio-environmental theories such as legislation, social action and intersectoral collaboration, in order to facilitate behaviour change through changes in health determinants and in increasing supportive environments for health. These two understandings of health promotion often run concurrently but ‘slippage’ to the individual behavioural focus is common (Baum, 2011).

As health promotion theory and practice evolved over time, so too has evaluation. This chapter has traced the development of evaluation theory and practice from the

1960s to the present day. In that time, evaluation development has introduced and promoted a move from a black box approach to a theory-based approach that attempts to unpack the intermediate steps and the mediators between inputs and outcomes. However, a range of evaluation approaches are currently in use and debate continues about the relative value of top-down, experimental and bottom-up context-specific methods (Patton 2011). Using a theory-based approach, instead of being focussed on choosing a method to demonstrate causation (with all the difficulty that brings in a community setting), the evaluator's task is to uncover the theory or program logic that underpins the steps within a program and then find appropriate ways to test the theory. The evaluator still brings particular expertise but engages with the program stakeholders to clarify their expectations and theories in order to make the evaluation findings more relevant and useful. The realist approach additionally emphasises the importance of including in the evaluation the context and mechanisms influencing behaviour change rather than trying to control these out.

Despite these developments in theory, and in the evaluation literature, much evaluation practice remains tied to a positivist paradigm. As Hancock suggests, researchers need to address the power structures that maintain 'the dominant research paradigm and processes of peer review, funding and publication' (Hancock, 1993 p24).

Community-based health promotion initiatives are often complex interventions in complex settings and as such, they are challenging to evaluate and need new ways of thinking. Evaluators are recognising the dynamic interactions and networks at work in a community-based intervention and are searching for evaluation approaches that can understand how these contribute to the achievement of outcomes. The next

chapter discusses how my published evaluations and the work leading to these reflect the struggle to undertake evaluations of complex initiatives within a predominantly positivist paradigm.

CHAPTER THREE: PUBLISHED PAPERS

3.1 Introduction

This chapter introduces my published work that forms part of this thesis. The full papers are included as an Appendix. Here, for each paper, the abstract is presented followed by a brief description of the study and its context. This draws on the literature review in Chapter Two and addresses Research Question One.

1. What was the historical health promotion and evaluation context for my publications and how did this influence my evaluation work?

The theoretical, methodological and practical challenges faced in conducting the research and evaluation are then analysed. These aspects are considered together as they are often intertwined. This addresses Research Questions Two and Three:

2. How do my publications reflect evaluation developments prior to 2008, including the changing role of the evaluator, in relation to community-based health promotion initiatives?

3. What are the contemporary theoretical, methodological and practical challenges in conducting community-based health promotion evaluations?

The final section of the chapter integrates the lessons from all the papers and describes the issues and questions that arise from the work and addresses Research Question Four:

4. What are the overall lessons from the evaluation practice presented in my publications and how do they inform new approaches to evaluation of community-based health promotion initiatives?

The timeline and major evaluation approach used for each piece of work leading to the published article is outlined in Table 3.1.

Table 3.1 Publication timeline and major evaluation approaches

Evaluation date	Evaluation approach	Journal article	Abbreviation used in text
2000 -2001	Process, impact, generative	Smith, A., Coveney, J., Carter, P., Jolley, G. and Laris, P. (2004) The Eat Well SA project: an evaluation-based case study in building capacity for promoting healthy eating. *Health Promotion International* 19 (3) 327-334.	*Eat Well SA*
2004 -2006	Document analysis	Baum, F. Jolley, G. Hicks, R. Saint, K. & Parker, S. (2006) What makes for sustainable Healthy Cities initiatives? - a review of the evidence from Noarlunga after 18 years, *Health Promotion International*, 21 (4) 259-265.	*What makes for sustainable Healthy Cities initiatives?*
2003 -2004	Meta-evaluation of qualitative evaluation reports	Jolley, G. Lawless, A. Baum, F. Hurley, C. and Fry, D. (2007) Building an evidence base for community health: a review of the quality of program evaluations. *Australian Health Review*, 31 (4) 603-610.	*Building an evidence base for community health*
2005 -2007	Resources for health promotion evaluation	Jolley G. Lawless A and Hurly C. (2008) Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion. *Health Promotion Journal of Australia,* 19 (2) 152-157.	*Framework and tools*
2004 -2006	Action research, case studies	Jolley G. (2008) Evaluation of an action research project in workforce development and organisational change: Healthy Ageing – Nutrition. *Evaluation Journal of Australasia*, 8 (1) 11-19.	*Healthy Ageing – Nutrition*

3.2 Eat Well SA Evaluation

Smith, A., Coveney, J., Carter, P., Jolley, G. and Laris, P. (2004) The Eat Well SA project: an evaluation-based case study in building capacity for promoting healthy eating. *Health Promotion International* 19 (3) 327-334.

Abstract

The term 'capacity building' is used in the health promotion literature to mean investing in communities, organizations and structures to enhance access to knowledge, skills and resources needed to conduct effective health programs. The Eat Well SA project aimed to increase consumption of healthy food by children, young people and their families in South Australia. The project evaluation demonstrated that awareness about healthy eating among stakeholders across a range of sectors, coalitions and partnerships to promote healthy eating and sustainable programs had been developed. The project achievements were analysed further using a capacity-building framework. This analysis showed that partnership development was a key strategy for success, leading to increased problem-solving capacity among key stakeholders and workers from education, child care, health, transport and food industry sectors. It was also a strategy that required concerted effort and review. New and ongoing programs were initiated and institutionalized within other sectors, notably the child care, vocational education and transport sectors. A model for planning and evaluating nutrition health promotion work is described.

3.2.1 Description and context

The study reports on the evaluation of the *Eat Well SA* project. This project aimed to increase the consumption of healthy food in the South Australian population and to facilitate strategies for doing this based on the Ottawa Charter for Health Promotion (World Health Organization, 1986a). The main approach of the project was to work to put food issues onto the agenda of other agencies by building intersectoral and interagency partnerships. The project used a capacity building approach with three domains: i) health infrastructure and service development; ii) problem solving; and iii) maintenance and sustainability. The aim of the publication was to describe the project, disseminate information about the methods and outcomes of the evaluation, and present a model for planning and evaluating capacity-building health promotion.

The *Eat Well SA* program management invited competitive tenders for the evaluation in November 1999. The evaluation brief required the successful consultants to evaluate key aspects of the program from its inception in October 1997 to October 2000, by developing an evaluation plan, undertaking data collection (including assisting project staff to develop data collection systems), managing data entry and analysis, producing progress reports and discussing findings and recommendations with staff and management groups, and writing a final evaluation report. A colleague (PL) and I were the evaluation consultants for the project, appointed in January 2000.

3.2.2 Theoretical, methodological and practical challenges

The evaluation in part took the standard path for this time of assessing process, particularly for the social marketing components of the project: Were the activities implemented as planned? Who was reached? What was the quality of the materials and components of the project? (Hawe, et al., 1990). Impact evaluation assessed the immediate effects of the program (Hawe, et al., 1990), in this case assessing changes in individual and organisational participants and policies. The more innovative generative evaluation component investigated the mediating role of the context in which the program was implemented and the interaction between stakeholders (Pawson & Tilley, 1997). Use of program logic models (Weiss, 1998) and realist evaluation (Pawson & Tilley, 1997) in the evaluation of public health programs was at an early stage when this evaluation took place in 2000/2001 and so this generative evaluation is an early example of assessing changes in relationships and health promotion capacity rather than the more traditional focus on individual behaviour change. An evaluation framework was developed based on Hawe and colleagues' work (1990) to illustrate the relationship between the various components of the evaluation.

This evaluation was my first practical experience of using a program logic model and the concept of generative evaluation. Previously, my evaluation frameworks were focused on process evaluation of strategies and short-term impact evaluation of objectives. I tended to consider process and impact evaluation separately, despite recognising that strategies had to be implemented as planned and this needed to be assessed before impact evaluation. This paper includes the program theory model developed for *Eat Well SA* which describes the logical links between the capacity building outcomes and impact indicators and the project strategies. Thus, the paper illustrates early use of logic modelling and program theory to develop indicators for evaluation.

In response to the multiple program components, the data collection was designed around the 14 sub-objectives/ strategies. The evaluators worked in consultation with the project managers to develop evaluation questions and data collection methods and to identify potential groups of respondents. Qualitative and quantitative data were analysed for their contribution to process, outcome or generative findings as appropriate.

Capacity building was identified by project management and the Evaluation Working Group as a key outcome of the project. The capacity building approach developed for the program has five sequential steps (see Figure 3.1). Each step has associated program activities that are the subject of the process evaluation and each step also guides the development of indicators and evaluation questions, for example:

> Indicator 1: Increased awareness among key stakeholders about food security and healthy eating.
>
> Activities: dissemination of information to key stakeholders through reports,

newsletters, forums, meetings, conference presentations

Process evaluation questions: what information was produced? Who was it disseminated to? Who was reached and who was missed?

Impact evaluation questions: How did awareness about food security and healthy eating change?

Generative evaluation questions: Have the relationships between organisations/agencies and other key stakeholders changed in relation to provision/promotion of healthy food? Is there a change in organisational understanding of the importance of action on food security and healthy eating?

Figure 3.1 Capacity building approach for Eat Well SA evaluation

The program logic model links the capacity building strategies of the project with generative change. The generative evaluation introduces the notion of sustainability and the evaluation discusses the benefit of the capacity building approach and the sustainability outcomes that follow. That is, while an evaluation may link modest health gains directly to the program, engagement with others leads to a greater potential to tackle health issues in the future. As evaluators, we argue in the paper

that a narrower focus on behaviour change would miss the important value-adding of capacity building. The paper demonstrates that capacity building is a valid and useful focus for health promotion evaluation in contrast to assessing only change in individual health-related behaviours.

The *Eat Well SA* project worked in partnership with many non-health sectors including schools and child care services, the food industry, non-government welfare organisations, and the environmental sector. The effectiveness of these partnership strategies was assessed in the evaluation. For example, effective partnerships were found with 50 organisations across all strategy areas of the project. Work with the child care sector and with the food freight transport industry to rural and remote areas of South Australia were particularly valuable in increasing awareness of healthy food and the need to improve food supply. Some partnerships were less productive and these seem to be related to different professional paradigms (e.g. community development approach versus professional nutrition) or different specialist knowledge and language required (e.g. with the environmental sector).

The evaluation did reveal that the intended community development approach of the small grants did not always lead to empowerment with the recipient groups concerned about the extensive monitoring and documentation required. There was also some tension between the community development intention and the nutrition education model which took precedence in some of the small grant cases.

As the evaluation report (Laris, Jolley, & Smith, 2001) and the paper conclude, the evaluation had to contend with a program that was diverse, often opportunistic in its activity focus and not fully articulated in terms of approach and rationale. A practical challenge was that when the evaluation commenced in January 2000, the program

had been running for over two years and goals and objectives were set. As evaluators we therefore had little opportunity to influence the goals and objectives of the program, nor, within the constraint of the consultancy, was there capacity to fully engage all stakeholders in the development of the evaluation framework and program logic. In part, this was due to the difficulty of engaging with the large community of stakeholders to the project and in part due to the established nature of the project development and implementation at the time that the evaluation commenced. Instead, a small Evaluation Working Group (including co-authors AS, JC and PC) was established and PL and I met regularly with this group.

This paper demonstrates that, within the limitations imposed by mainstream positivist thinking, a less linear evaluation drawing on stakeholder perspectives to identify underlying program theory was able to be undertaken. Thus, for the *Eat Well SA* evaluation, a comprehensive framework of key questions was developed to assess i) process (what happened, who was reached and what methods were effective) ii) impact (what changes were observed in knowledge awareness behaviour or policy) and iii) generative (what changes in organisational relationships and the healthy eating context).

The evaluation report notes that direct data on the achievement of the goal (to increase the consumption of healthy food by young families) and objectives was not available or feasible to obtain. Instead, drawing on the program logic model, intermediate outcomes such as changes in range and quality of food available (for example, in child care centres), awareness and beliefs about healthy food (for example, in child care centre staff and managers; school communities; families and those who work with them; small grant recipients), and organisational policy changes

(for example, small grant recipient organisations) were assessed. How to evaluate changes in longer term health outcomes and attribution of these to the intervention was a remaining question.

3.3 What makes for sustainable Healthy Cities initiatives?

Baum, F. Jolley, G. Hicks, R. Saint, K. & Parker, S. (2006) What makes for sustainable Healthy Cities initiatives? - a review of the evidence from Noarlunga after 18 years, *Health Promotion International*, 21 4: 259-265

Abstract

This paper examines the factors that have enabled the Healthy Cities Noarlunga (HCN) initiative to be sustainable over 18 years (1987–2005). Sustainability related to the ability of the initiative to continue to operate continuously in a manner that indicated its existence was accorded value by the community and local service providers. The analysis is based on a narrative review of 29 documents related to HCN, including a number of evaluations. Nine factors emerged as important to ensuring sustainability: strong social health vision; inspirational leadership; a model that can adapt to local conditions; ability to juggle competing demands; strongly supported community involvement that represents genuine engagement; recognition by a broad range of players that Healthy Cities is a relatively neutral space in which to achieve goals; effective and sustainable links with a local university; an outward focus open to international links and outside perspectives; and, most crucial, the initiative makes the transition from a project to an approach and a way of working. These sustainability factors are likely to be relevant to a range of complex, community-based initiatives.

3.3.1 Description and context

The aim of this study was to identify the factors that have contributed to the sustainability of Healthy Cities Noarlunga (now Healthy Cities Onkaparinga) over some 18 years, despite political, structural and organisational change over this time. The study was jointly driven by the research team (FB and myself), and by other Healthy Cities members, prompted by a long association between SACHRU and Healthy Cities Noarlunga, and a desire to investigate the reasons for the long and ongoing life of the initiative. One of the co-authors, Richard Hicks, was the Chair of Healthy Cities Noarlunga at the time this paper was written and had held this

position since the inception of the initiative in 1987. Fran Baum was also pivotal to the establishment of Healthy Cities Noarlunga and, through SACHRU, has maintained an association providing academic support and links to the international Healthy Cities movement. The final two co-authors were also long-time active members.

The paper is based on a review of Healthy Cities Noarlunga, one the original pilot Healthy Cities in Australia and still active in 2012 as Healthy Cities Onkaparinga. A brief history of Healthy Cities Noarlunga is given and a review of why evaluation of this type of complex community-based initiative is difficult. A narrative review of Healthy Cities Noarlunga documentation was undertaken and from this, and a review of relevant Healthy Cities literature, nine sustainability factors were identified. The paper concludes that the sustained effort of Healthy Cities Noarlunga is likely to have contributed to improved health, thus suggesting plausible links between program activities and short and long term outcomes.

3.3.2 Theoretical, methodological and practical challenges

Sustainability is a key criterion for success in health promotion programs (Pluye, Potvin, Denis, & Pelletier, 2004; Shediac-Rizkallah & Bone, 1998) and therefore how sustainability is conceptualised and evaluated are critical questions. This study used document analysis to develop a framework of sustainability factors for a community-based health promotion program. The meta-evaluation method drew on 29 documents published since 1987. Most of these documents would be classified as ‘grey literature’ comprising annual reports, activity reports, small scale evaluations and other reports. The authors’ knowledge of the Healthy Cities Noarlunga initiative was critical both in identifying and accessing documents and in interpreting the content. The dearth of formal evaluations of Healthy Cities has been noted elsewhere

(Neiman & Hall, 2007; Poland, 1996a) and reflects the challenges in evaluating community-based initiatives and the lack of adequate resources to develop and conduct appropriate methods. In this case, I was able to draw on my co-authors' expertise to provide a mix of theory and practice knowledge to assist with the analysis and interpretation of the documentation pertinent to the sustainability of Healthy Cities Noarlunga. Thus, the study was informed by evidence from academic literature and knowledge from those involved in the initiative.

From the document analysis and the researchers' (FB and GJ) intimate understanding of the project, nine factors were initially identified to build up a conceptual model of sustainability. The Healthy Cities Noarlunga member co-authors (RH, KS and SP) verified the importance of these factors and provided examples of how they had worked in practice. In the paper, each of the factors is supported by evidence from the academic literature or from the Healthy Cities Noarlunga documents.

Reflecting back on how sustainability was analysed in the study suggests an intuitive context-mechanism-outcome construction as described in section 2.4.4. Thus, the nine factors could be reconstructed as context and mechanisms of a realist evaluation (Pawson & Tilley, 1997), leading to the outcome of sustainability. For example, Healthy Cities Noarlunga was established at a time when the South Australian State Government was strongly supportive of social health and primary health care policy designed to address health inequities. This meant that the context was right for the development of a social health vision that was the basis of Healthy Cities Noarlunga. The document analysis identified the consistent and committed leadership in Healthy Cities Noarlunga as an important mediator or mechanism in maintaining enthusiasm for working towards achievement of the vision.

The *What makes for sustainable Healthy Cities initiatives?* paper reports that partnerships were important factors in sustainability: partnerships with community, local small business, local government, the research community and international Healthy Cities links. The evaluation suggests that these partnerships all helped to ground the initiative in local issues, making it relevant for the local context. Partnerships were also able to provide credibility and support for a community-based group with very limited resources.

So, this study used program and evaluation documentation, and evaluator and stakeholder expertise, to identify critical factors (context and mechanisms) leading to sustainability. The paper demonstrates the enabling factors to sustaining community-based health promotion initiatives over the long-term, which Shediac-Rizkallah & Bone, 1998 note is critical if such projects are to bring significant benefit to the community.

The Healthy Cities Noarlunga initiative was shown to lead to change that is likely to lead to health improvement in the longer term. For example, some community members reported that their participation had led to lasting positive impacts such as increased supportive networks and skill development. Wider, population level changes were also reported, such as stronger networks between participating agencies and community groups for environmental advocacy, and small business workplace safety programs. Examples of changes achieved that have a direct link to health outcomes include reducing pollution in the local estuary and removal of injury hazards in the community. The conclusion is that health promotion initiatives such as Healthy Cities are more likely to be able to show successful outcomes if they are maintained over the longer term. The question of how discreet health outcomes can

be directly attributed to health promotion activities remains. One approach is to draw on other evidence to suggest that the changes seen are likely to lead to longer term health gains (Baum, Jolley, & Bament, 2001). This, of course, is the essence of theory-based evaluation that makes clear the logical links between steps in the program.

3.4 Building an evidence base for community health

Jolley, G. Lawless, A. Baum, F. Hurley, C. and Fry, D. (2007) Building an evidence base for community health: a review of the quality of program evaluations. *Australian Health Review*, 31 (4) 603-610.

Abstract

An assessment of the quality of program evaluations conducted in South Australian community health services investigated how effective evaluation reporting is in producing an evidence base for community health. Evaluation reports were assessed by a team of reviewers. Practitioner workshops allowed an understanding of the uses of evaluation and what promotes or acts as a barrier to undertaking evaluations. Community health services do undertake a good deal of evaluation. However, reports were not generally explicit in dealing with the principles that underpin community health. Few engaged with program theory or rationale. Typically, reports were of short-term projects with uncertain futures so there may seem little point in considering issues of long-term health outcomes and transferability to other settings. The most important issue from our study is the lack of investment in applied health services research of the sort that will be required to produce the evidence for practice that policy makers desire. The current lack of evidence for community health reflects failure of the system to invest in research and evaluation that is adequately resourced and designed for complex community settings.

3.4.1 Description and context

The aim of this research was to assess the quality of evaluations of programs run in five Community Health Services in the metropolitan region of South Australia. The paper presents the results of that review. The research began in 2003 and undertook a review of program evaluations conducted during the three years, 1999 – 2002. At that time, the state-funded and managed Community Health Services were one of the

main providers of non-medical primary health care and health promotion and shared the philosophy of the Ottawa Charter. Multi-disciplinary staff undertook health promotion activities including screening, individual and group health education, community development, advocacy and contributing to public health planning (Baum, 2002).

The idea for the research arose from managers of the services who wanted to respond to pressure from funders to demonstrate effectiveness. In a context of evidence-based medicine and health services provided in a competitive funding context, they believed that not having evidence of effectiveness could disadvantage community health services and their health promotion activities. The research also built on some earlier work by SACHRU researchers, including myself, (Baum, Duffy, & Jolley, 2003) that examined four ways in which community health evidence of effectiveness could be judged: economic evaluation; use of services' routine data bases; systematic reviews; and performance indicators. That report concluded that a more systematic approach was needed to assess the effectiveness of community health services

3.4.2 Theoretical, methodological and practical challenges

Evaluation is crucial for developing and maintaining effective and equitable health promotion activity (Baum, 2003; Green & Tones, 2010; Kelly, et al., 2007; Poland, 1996a). Lack of evidence of effectiveness can disadvantage community health and primary health services undertaking health promotion activities when funding allocation decisions are made.

As part of SACHRU's commitment to participatory research and evaluation, a reference group that included community health practitioners from the services was set up to guide and support the study. SACHRU had an established relationship with

the services and it was important to maintain these relationships as the researchers were dependent on the services for access to the evaluation reports. A practitioner from each service joined the academic researchers on the review team to finalise the review questions and undertake the assessment of evaluation reports. Training and support was provided to all reviewers to ensure consistency in the assessment process. This use of practitioners on the review team helped to establish a sense of ownership in the research for services and reduced the potential for the exercise to be seen as threatening. It also ensured that academic and practice wisdom was complementary in the study.

The planned project remained flexible so that, for example, evaluation workshops were conducted at each of the six services at their request. The process varied but in general participants were invited to discuss current uses of evaluation, and the enablers and barriers for them in conducting and using evaluation. Thus, the already established trust and valued relationship between the evaluation team at SACHRU and the services was a critical factor in the success of the project's implementation. However, the workshops were not planned or budgeted for at the beginning of the project but were met out of existing resources as it was believed important to support ongoing stakeholder engagement.

From a collection of all documented evaluations identified by the five services (n=120), 93 met the inclusion criteria and were included in the review. The majority of the evaluations were conducted and reported on by the practitioner(s) involved in running the program. Since all the reports contained at least some qualitative data (and indeed many were almost entirely based on qualitative data) the researchers developed a review framework for systematic review of qualitative program reports

assessing planning, program logic and evaluation. Criteria for assessing the quality of qualitative evaluation research were not well-developed at this time. We developed an assessment framework based on the work of Popay and colleagues (1998) and Rychetnik and Frommer (2001), and adapted this to ensure the inclusion of health promotion and primary health care principles including community participation, equity and recognition of the social determinants of health.

The evaluated programs were categorised as individual services, group programs or community development. About 43% were judged to be community development type initiatives and about half of these ran for more than 12 months. Thus, only about one quarter of the evaluations were of community-based health promotion initiatives as defined in this thesis but the total represents the reported evaluation activity of the services at the time. A comparison with service activity shows that, while individual services made up 47% of activity by time (Department of Human Services, 2001), these delivered only 3% of evaluation reports for our review. Community development and health promotion initiatives made up 45% of service time and 43% of the evaluation reports. Thus, it appears that one-to-one services are not evaluated to the same extent as other activities or are evaluated in a way that did not result in a report meeting the criteria for our study.

The review found that most evaluation in the context of these primary health care services consisted of practitioner-led internal quality control reports for small scale programs. The assessment showed there was little application of theory or articulation of an underlying rationale for programs, suggesting that, although these concepts were quite well-established at this time in the evaluation community (Coryn, Noakes, Westine, & Schröter, 2011) they had not become part of health

services thinking. Equity, community participation and intersectoral collaboration, important principles of health promotion, were seldom assessed in the reviewed evaluation reports.

It appeared that limited use was made of the evaluation reports inside or outside of the service. Where reports were used in the service it was mainly for purposes of planning and review of process rather than assessing effectiveness of a program. It seems that, despite the services recognising the importance of evaluation in providing evidence for primary health care and health promotion, there was a lack of skill base and resources to implement evaluation beyond basic process issues. Further, services were unclear how to make best use of evaluations in order to improve their activities or contribute to the evidence-base for their activities. Similarly, South and Tilford (2000) found health promotion specialists working in the United Kingdom undertook a range of evaluation activities but it was uncommon to find sufficient extensiveness or methodological rigour to contribute significantly to the evidence base for health promotion.

This research is important because it highlights the enhancers and barriers to practitioner-led evaluation in health promotion programs. The paper revealed the lack of theory or articulation of the underlying rationale for many health promotion programs, and confirmed the associated need for training and resources in evaluation. Practitioner respondents in the study identified the need for skill development and training and a ‘culture of evaluation’ in order to promote quality evaluation. Lack of time and resources were cited as the main barriers.

From my perspective as an evaluator, this paper illustrates the conflict between trying to develop overarching and generalisable findings from health promotion

evaluations and being aware of the local context that influences a program's design and implementation. The paper lends support to the notion that a comprehensive theory-based evaluation is difficult in small-scale and resource-limited project evaluation of the type commonly conducted by practitioners for community-based health promotion initiatives. The evidence from this paper suggests that while a great deal of (questionable quality) evaluation is conducted in the services in this study, it generally did not articulate or make use of any underlying theory or rationale. Furthermore, the application of major health promotion principles is seldom included as a focus of the evaluation.

Since it seems unrealistic to expect health promotion practitioners to have the necessary skills and resources to undertake a theory-based evaluation without support, one role for an evaluator would be to work with individual programs to construct a theory and use this in the evaluation. Programs seldom have sufficient resources to engage evaluation consultants (even if external evaluation is considered desirable) so acting as mentor and support may be a better way to build an evaluation culture within services. This approach also means that cumulative learning over many initiatives is unlikely to occur if each is evaluated independently. An alternative would be for an institution to be resourced to undertake collaborative evaluation with health promotion practitioners and services. This is similar to the SACHRU model which is able to provide some direct support in this way, to provide workforce development in research and evaluation, and to undertake dissemination of accumulated experience. In my experience of more than 15 years, this model does appear to have contributed to an enhanced culture of evaluation in South Australian primary health care services and the large number of evaluation reports identified in this research supports this view. Another model would be to focus supported

evaluation effort on programs that have more resources, are larger or particularly innovative, in order to maximise learning about effectiveness. Practitioners could then undertake formative and process evaluations on smaller scale programs where longer term outcomes are too difficult to measure and attribute. Nutbeam (International Union for Health Promotion and Education, 2000) for example, considers that health promotion programs are only worth evaluating for effectiveness if they have a reasonable chance of success. This means they are:

- Planned on the basis of an assessment of evidence from epidemiological, behavioural and social research which indicates reasonable linkages between short medium and long-term outcomes
- Informed by established theory
- Implemented under the necessary conditions for success (public and political awareness of issue, capacity for program delivery, resources to implement and sustain)
- Of sufficient size, duration and sophistication to detect outcomes above the background noise (International Union for Health Promotion and Education, 2000 p10).

The study revealed much information about the current state of evaluation in South Australian community health services. While we did not intend to draw generalisable conclusions as might be found in a systematic review, five good practice case studies were described in the full report (Jolley, Baum, Hurley, & Fry, 2004). SACHRU used this study to make a number of recommendations about ways to increase the capacity of services to undertake evaluations. Specifically, with colleagues at SACHRU, I contributed to the development of two resources for health promotion practitioners: an evaluation and reporting template (South Australian Community

Health Research Unit, undated) and a framework to assist with evaluation of community participation, partnerships and equity. The latter resource is described in my next publication: *Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion.*

3.5 Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion

Jolley G. Lawless A and Hurly C. (2008) Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion. *Health Promotion Journal of Australia*, 19 (2) 152-157.

Abstract

Issue addressed: This paper reports on the development of a planning and evaluation framework and tools to assess key principles of primary health care/health promotion: community participation, collaborative partnerships and a focus on equity. The focus of the tools is on planning and process evaluation with some outcome questions included.

Methods: Following a scan of literature, the framework and tools for each component were developed. The tools were road-tested with colleagues and trialled by workshop participants.

Results: A framework and tools for each of the components and ways to assess how effectively they are applied at the program and practice level was developed. The tools attempt to deal with evaluation challenges by providing primary health care/health promotion practitioners and evaluators with a framework to examine these components of their work.

Conclusions: Planning and evaluation are regarded as routine in good practice. As health promotion practice and programs are shaped by principles such as partnerships, participation and equity, it is important that we also apply an evaluation lens to these components. Sound planning and evaluation allows practitioners to explain how and why these principles are integrated into their work and what is achieved.

3.5.1 Description and context

The aim of this paper was to disseminate a planning and evaluation framework for three important aspects of health promotion practice: working in partnership,

community participation and a focus on equity. This work follows on from the '*Building an evidence base for community health: a review of the quality of program evaluations'* paper. With the two co-authors, I designed a framework and tools to assess community participation, collaborative partnerships and equity in small scale health promotion initiatives. My particular focus in the paper was on the community participation component. The development of the framework was a collaborative process, drawing on academic and practice wisdom, using field trials with practitioners and a workshop presented at a health promotion conference (Jolley, Lawless, Hurley, Biedrzycki, & Ramanathan, 2007).

The framework and tools for evaluation of community participation (developed by GJ) discusses levels and stages of participation, and the benefits of involving community members and consumers in health care decisions. The tool covers four dimensions: i) extent and scope of participation; ii) working together; iii) capacity and support; iv) impacts of participation.

The partnerships tool focuses on evaluation in an unpredictable and changing, context. It suggests three dimensions for evaluation: i) partnership effectiveness; ii) benchmarking/describing current status; iii) targeting strengths and weaknesses for further development. A reflexive process rather than solely assessment of goal achievement is recommended.

The equity component considers fairness and the opportunity to attain full health potential. Three dimensions are included in the tool: i) equal access to available care for equal need; ii) equal utilisation of available care for equal need; iii) equal quality of care for all. Taken together these should act to reduce inequities in the outcome of care. These dimensions can go beyond provision of health care to include those

interventions that aim to influence the social and economic factors largely outside the health sector that mediate or influence social and economic disadvantage. In this way, access, utilisation and quality can be thought of in regard to 'resources for health' and equity principles applied to broader health promotion interventions. This helps to clarify the ways that achievement of equity can be evaluated in health promotion programs.

3.5.2 Theoretical, methodological and practical challenges

The previous paper *'Building an evidence base for community health'* demonstrated the lack of evaluation of community participation, collaborative partnerships and equity in the small scale programs reviewed in the study. This paper offers a framework and tools for practitioners to use in planning and evaluation of health promotion programs, based on a theoretical understanding of community participation, collaborative partnerships and equity, and is aimed at assisting practitioners to build an evidence base for their practice and for health promotion more generally.

By focussing on community participation, partnerships and equity, the process draws together evaluation theory and practice with health promotion principles. The paper also considers complex environments of frequent change and multiple stakeholders. For example, the purpose and aims of a partnership may change over time and people or organisations may join or leave, in response to contextual developments. This means the assessment tools will need to be flexible and re-applied following change.

For the community participation evaluation tool, community participation theory is drawn on to construct the evaluation framework. This theory is based on community participation as an ethical and democratic right and the notion that involving people

in health care decision making lead to changes that will improve health (Oakley & Kahassay, 1999; Rifkin, 1996; South Australian Community Health Research Unit, 2000). Four dimensions for evaluation of community participation are identified: the extent and scope of participation (numbers and characteristics of people); the range of processes for working together; the extent of capacity and support for service providers and community members to work in partnership; and, finally, the impacts of participation and changes made as a result of community input. Community participation is a complex and contested component of health promotion practice and the paper concludes that sound evaluation practice can lead to building the evidence base for the value of this approach in terms of strengthening rights and improving health outcomes. An evaluation can provide information to clarify how and why the principle of community participation is integrated into programs.

Evaluation is expected as part of good practice, contributing to our knowledge of what works, for whom and why. However, as we have seen, practitioners are seldom resourced sufficiently to undertake systematic evaluation. Achievement of health promotion principles of community participation, partnership and equity are not easy to evaluate due to these being contested terms, conceptually difficult to understand, grounded in value judgements and reflecting the complexity of interaction with the setting or context. Nonetheless, if these are accepted as critical components of health promotion initiatives, it is important to include these aspects in evaluation (Funnell & Rogers, 2011; Wass, 2000).

One of the challenges that arose in workshopping the evaluation framework was the varied and diverse understanding of health promotion concepts held by practitioners. Much time was spent in discussing case study examples and coming to an agreement

about terms. This lack of consensus across services and practitioners seems inevitable given the multi-disciplinary workforce and the diversity of organisational structure of primary health care and health promotion services. It is therefore difficult to develop and establish consistent ways to assess attributes such as community participation, partnerships and equity across this broad range of activities and structures.

A further challenge is that health promotion practitioners are likely to prioritise practice over planning and evaluation, and, for many, demand for direct service provision limits capacity for health promotion activity. This raises two overlapping questions: i) what should be the role and responsibility of health promotion practitioners in the evaluation of programs and ii) to what extent are practitioners able to engage in collaborative evaluation with external evaluators? There is a potential conflict between a collaborative evaluation approach and the capacity of practitioners and other stakeholders to be involved in evaluation. Further, what is the role of the evaluator in advocating for, and assisting with development of, an evaluation culture? The high quantity of evaluation reports identified in '*Building an evidence base for community health'* suggests that, over the years, SACHRU may have contributed to the building of an evaluation culture in South Australian primary health care services.

3.6 Evaluation of an action research project in workforce development and organisational change: Healthy Ageing – Nutrition

Jolley G. (2008) Evaluation of an action research project in workforce development and organisational change: Healthy Ageing – Nutrition. *Evaluation Journal of Australasia*, 8 (1) 11-19.

Abstract

This article reports on the evaluation of an action research project designed to support workforce development in the promotion of healthy nutrition for older people. The evaluation methodology was grounded by the action research approach of the project and focused on case studies of the 10 partner organisations. Findings indicate that the Healthy Ageing-Nutrition Project has resulted in a large increase in awareness and knowledge about healthy ageing and nutrition in the case study organisations, and to a lesser extent, in the broader health and aged care sectors. For the case study organisations it seems likely that transformational change has been made through the project's work of building capacity, mediating and facilitating change and providing resources. Support at board and management level, as well as thoughtful development of the workforce, were critical success factors in bringing about organisational change. The main challenge was identified as time and resources needed. Follow-up evaluation of the health outcomes from nutritional assessment, screening and intervention should also be implemented in order to provide further evidence of the value of this effort.

3.6.1 Description and context

Healthy Ageing – Nutrition was an action research project aiming to improve nutrition capacity in the aged care workforce and community-based organisations providing aged care services. Over two years, from March 2004 to March 2006, the project worked with ten organisations to develop action plans, provide resources and other assistance and bring the participants together for shared learning. Although the stated target was workforce, this was defined very broadly to include volunteers and carers. Each organisation's action plan was developed individually but based on a proforma shown in Table 3.2. The evaluation reflected the action research approach, using a mix of internal (project manager) and external (GJ) evaluators. The external evaluation reported in this paper aimed to assess the strength of the action research process and progress towards the common project objectives. The external evaluation contract began in August 2004 early in the life of the project. The evaluation was consultative throughout, working closely with the project manager and also with the project advisory group. A qualitative case study approach was used to evaluate the participating organisations' progress in achieving their action plans.

Table 3.2 Healthy Ageing – Nutrition Action Plan Proforma

Achievement	Task/Action	Measurement	Who/when	Capacity Development
We hope to achieve:	In order to do this we will:	We'll know when this is achieved because: We'll reflect on this by:		This change is: • Policy development • Organisational development • Workforce development • Resource allocation • Leadership • Partnerships

3.6.2 Theoretical, methodological and practical challenges

This paper provides an example of the challenges of evaluation of a health promotion initiative in a community setting. The evaluation is a series of case studies based on qualitative data from the stakeholder organisations. The evaluation was able to establish a model of enhancers and barriers to organisational change based on the ten case studies.

Data collection was in the form of two interviews with each participating organisation leader, a series of interviews with the project manager, and a review of the action plans. Evaluation was both formative and summative. The first two interviews investigated how the organisation was 'preparing for the journey' and how things were unfolding 'on the road'. These formative evaluation findings were fed back to the organisations. For the summative evaluation, action plans were analysed to check reported achievements, as identified in the final interview, against what had been planned. Most reported achievements centred on increasing awareness and knowledge about nutrition and older people, or the introduction of new policies,

screening procedures or menus. It was beyond the scope of the evaluation to measure changes directly but some documentation supported the self-reports. Perceived enablers and barriers to change were also collected as ‘reflections on the journey’. The project manager and stakeholder organisations were kept informed of evaluation findings and participating organisations had opportunities to comment on their own case study.

The ten stakeholder organisations each developed their own action plan (similar to a simple program logic model) under the guidance of the overarching project. The internal and external evaluators worked with the participating organisations to evaluate their action plans and feedback was provided as the program progressed. This meant that the evaluation had to be sufficiently flexible to be implemented in different organisations, and respond to their needs and capabilities, while aiming for some consistency in order to derive general lessons.

The participating organisations developed their action plans differently. Most needed a lot of help, particularly at the beginning, and in fact preferred the project manager to develop the action plan for them initially. Thus, the capacity of the organisations, or the person whose role is was to undertake this activity, appeared limited and the evaluation suggests that a high level of support would be needed for most community-based organisations to undertake similar, but more complex, program logic modelling.

As the external evaluator, I argue in the evaluation report that the action research approach is appropriate for a project aiming to bring about organisational change in a multidisciplinary, multi-sector setting, using a participatory approach. Action research acknowledges the importance of practice wisdom as a way of developing

theory when other evidence may be hard to find or demonstrate. However, barriers to participatory evaluation were reported. Eight of the ten organisations reported that the main barrier to beginning action was a lack of time and resources within their organisation and the increase in work load due to participation in the project on top of usual duties (Government of South Australia, 2006). Engagement in the evaluation would therefore likely be seen as an additional burden and it is a testament to the project manager's skills and good relationships that all the organisations did in fact agree to two rounds of interviews and also attended the final feedback forum. On the other hand, project advisory group members described their role as information sharing, with only one respondent agreeing that the group had any influence on how the project was implemented. Given that this group was one of the main ways for stakeholders to be fully engaged in the action research process, this finding is disappointing. So, while some members of the group stated that they were able to influence how the project unfolded in their own organisation this did not extend to having an impact on how the overall project was implemented.

Strengths of the evaluation were that it was planned early in the life of project, engaged with stakeholders in its design and used a diversity of data collection methods and sources. A case study methodology was able to capture information from a range of stakeholders in each organisation and bring these data together to present some common themes in the form of a conceptual model of enhancers and barriers to organisational change. Thus, in terms of organisational change, the evaluation was able to draw lessons about 'what works for whom under what circumstances'. For example, the *Healthy Ageing – Nutrition* paper and the evaluation report (Government of South Australia, 2006) state that the smaller organisations, with less bureaucracy, seemed able to take on change more quickly

and one person acting as a change agent can have considerable influence. Larger organisations, and those with a large volunteer workforce, appeared to find it more difficult to shift focus or take on a new way of doing things. On the other hand, once change is accomplished in a larger organisation it seems to be more likely to be embedded and sustainable. This suggests that change management approaches may be more effective if they are tailored to the characteristics of the organisation.

In common with most health promotion initiatives, *Healthy Ageing – Nutrition* funding was time-limited (two years) and so longer term assessment of the sustainability of the changes was not possible. Neither was there an opportunity to follow up on potential health outcomes from the nutrition interventions put into practice. However, a theory of change model for the program, based on capacity building was able to describe the logic of the links between activities and immediate outcomes and the evidence that they are likely to lead to longer term outcomes. The evaluation also illustrates the importance of monitoring for unpredicted events, for example, an unexpected finding was that the program led to links between agencies working in nutrition and falls prevention.

3.7 Key lessons and implications for evaluation

This chapter has presented the aims, context and challenges for the research and evaluation practice described in my papers. This next section draws together the theory and practice described above and discusses two critical issues for the evaluation of complex community-based health promotion initiatives that arise. These issues are: i) theoretical and methodological issues, ii) practical and resource issues.

3.7.1 Theoretical and methodological issues

My papers illustrate some of the theoretical and methodological issues in evaluating community-based health promotion initiatives. Theoretical underpinnings of the programs evaluated in the papers were generally weak. Although there are numerous health promotion theories to draw upon (see, for example, Nutbeam, et al., 2010) as outlined in Table 2.1, explicit use of theory to design health promotion programs of the type described in my publications was rare. Rather, my evaluation role included negotiating with stakeholders to draw out their understandings and expectations of the program in order to make clearer the implicit theories and program logic. Using Nutbeams's classification, Table 3.3 categorises the main theories and models that underpin the programs that are the subjects of the papers.

Table 3.3 Theories and models for health promotion programs

Health promotion program	**Theory or model**
Eat Well SA	social marketing theory and intersectoral action models
What makes for sustainable Healthy Cities initiatives?	community mobilisation and empowerment theories, intersectoral action models healthy public policy framework
Building an evidence base for community health	many programs were concerned with social cognitive theory and/or health literacy models
Healthy Ageing – Nutrition	organisational change theory

These theories and models explain how the intervention is predicted to work and so assist in developing a program logic model and evaluation framework to identify key evaluation foci and questions (Bickman, 2000; Chen, 1990). Using program theory is likely to be relatively simple when the program is straightforward and, according the

stakeholders, based on an accepted theory and well-tested assumptions. However, for many community-based health promotion programs, the complex setting and multiple stakeholders mean that theories are contested and assumptions are un-tested (Birkmayer & Weiss, 2000; Mark, 2003). In this case, an evaluation design grounded in a specific program theory is likely to be limiting and may not capture key aspects of the program as it unfolds during implementation. For example, using only organisational change theory in the evaluation of *Healthy Ageing – Nutrition* may have missed the importance of the South Australian context and the intersectoral relationships that were formed. So several theories may be applicable to how the program is expected to work and the evaluation should be open to these. In other words, there may be several (contested) program theories operating at the same time (Birkmayer & Weiss, 2000; Mark, 2003) so the evaluation needs to test what actually happens against what was predicted by the program theory or theories (Patton, 1997) and potentially suggest new or different program theories . Further, the program logic model is not linear but may need to incorporate feedback loops as the context for the program changes (Poland, 1996a).

Within the constraints of conducting contracted evaluations, the evaluations reported on in my publications used methodologies that were appropriate for the initiative, reflected my expertise and values, and the value-base of SACHRU as my academic institution. None of the evaluations used experimental designs since establishing controls was not considered feasible or appropriate by the evaluators and the evaluation contractors accepted this view. Instead more interpretive approaches were used, mainly drawing on interview data to uncover stakeholder and participant perceptions of program quality and achievement. While an interpretive approach was able to deal in part with complexity and environmental influences on the program, it

limited the claims that could be made about attribution and causation, as discussed in Chapter Two.

Qualitative methods for research have become steadily more accepted over the last four decades but uptake has been slow in the field of evaluation (Evans, et al., 2007; Weiss, 1998). The advent of theory-based evaluation that is not tied to a particular methodology, encourages a choice to be made that best complements the evaluation purpose (Chen, 1990). This suggests that policy makers and program funders looking for evidence of effectiveness of community-based health promotion need to become more open to methodological diversity and the recognition that linear cause and effect can seldom be established in complex environments.

Another issue for evaluation is to develop methods to evaluate health promotion principles as identified in Chapter Two. Some methods to evaluate community participation, partnerships and equity are illustrated in my publications and these are discussed below.

Community participation has been defined as:

> 'The involvement of consumers in the development of health services. This can include involvement in policy development, strategic planning, service planning, service delivery and evaluation and monitoring' (Consumer Focus Collaboration, 2000).

Participation can be empowering through developing the skills and networks of participants and providing feedback to increase control over the program (Rootman, Goodstadt, Potvin, & Springett, 2001).

Participatory programs and their evaluations are likely to undergo developmental change over time as stakeholder views influence the program implementation and the context changes. These adaptations and changes makes it difficult to specify in

advance to funders the activities and expected outcomes (Guba & Lincoln, 1989). Unpredictable interactions and feedback loops, and changes to the program in response to stakeholder input, mean that the evaluation needs to be adaptive in order to capture these influences and their impact.

A framework to guide the evaluation of community participation is presented in *Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion*. While measuring opportunities for participation and numbers of participants is relatively simple, assessing the impact is more complex. There is good evidence to suggest that community participation is beneficial for participants and services (Oakley & Kahassay, 1999; Rifkin, 1996; South Australian Community Health Research Unit, 2000) and health promotion programs usually aspire to some level of participation. Even so, the rationale and mechanisms for community participation may not be spelt out so that this principle may not feature in the program logic or the evaluation. Table 3.4 illustrates how my evaluation work and publications have contributed to methods for evaluation of community participation.

Table 3.4 Evaluation of community participation

Health promotion program	**Evaluation findings**	**How evaluated**
Eat Well SA	Stakeholder rather than community participation Some participation from grant recipients	Stakeholder interviews Grant application process and feedback
What makes for sustainable Healthy Cities initiatives?	Community participation identified as one of the sustainability factors e.g. majority community membership on management committee	Document analysis
Building an evidence base for community health	Reported input to planning and implementation of programs limited	Review of written program evaluation reports
Healthy Ageing – Nutrition	Community participation limited to volunteers working in participating organisations High involvement from participating organisations in own case studies Limited in overall program	Case study interviews Program manager interviews

The evaluations were mostly limited to measures of the extent of participation, although the *What makes for sustainable Healthy Cities initiatives?* findings link the outcome of sustainability of the initiative to community participation. For the other evaluations, community participation was not a key focus of the programs and indeed *Building an evidence base for community health* identified this as a gap in the

evaluation reports reviewed for that study. Challenges include the profusion of terms and understandings related to community participation as a concept (Green & Tones, 2010) and the difficulty of identifying the links between community input and changes to the outcomes of programs. It would seem that community participation is not valued sufficiently by health promotion funders or program designers to be a focus of evaluation. Health promotion activists and evaluators may need to strengthen this aspect of their practice to ensure that community participation principles are adequately reflected in programs and initiatives.

Health promotion practice often requires working in partnership with other parts of the health system and with other sectors outside of heath. The principle of intersectoral collaboration recognises that achieving population health goes beyond provision of health services and encompasses efforts to address the social determinants of health. To do this, other sectors are required to recognise their role in contributing to population health. The engagement of partners or other sectors in health promotion initiatives is therefore often a topic for evaluation.

The evaluations reported on in my publications (Table 3.5) suggest the value of a partnership approach in tackling issues that go beyond health care to addressing the social determinants of health.

Table 3.5 Evaluation of partnerships

Health promotion program	**Evaluation findings**	**How evaluated**
Eat Well SA	Program worked with over 50 health and non-health partners, mostly effective and beneficial to achieving aims	Document analysis Interviews
What makes for sustainable Healthy Cities initiatives?	Management group is Intersectoral including community members and health/welfare and non-health agencies	Document analysis
Building an evidence base for community health	Partnerships with non-health agencies and non-government sector reported	Review of written program evaluation reports
Healthy Ageing – Nutrition	Project manager worked in partnership with participating organisations Limited evidence of new or strengthened partnerships	Case study interviews Program manager interviews

Challenges identified by the evaluations include potential tensions when organisations and sectors with different value-bases work together, and the need to develop a shared language and understanding of the issues. Evaluation of partnerships and intersectoral action is important in order to increase understanding of the costs and benefits of this approach. So partnership evaluation needs to ask

whether the right partners have been engaged, whether critical players have been missed and whether more has been achieved than would have been without the partnership and at what cost to both sides. Moreover, partnerships will not be static over the duration of the initiative but are likely to reflect a dynamic process as issues emerge that require new partnerships, and as players enter or leave the implementation space.

Equity is described as the fair opportunity to attain full health potential (World Health Organization, 1986b). As an important principle in health promotion, programs should not contribute to inequity – 'differences which are unnecessary and avoidable but, in addition, are also considered unfair and unjust' (Whitehead, 1991). Programs may be targeted at disadvantaged communities or may be universal. In either case, if equity is accepted as an underpinning principle of health promotion then it follows that evaluation design needs to consider any intended and unintended effects on equity in terms of the distribution of costs and benefits across stakeholder groups (Poland, Krupa, & McCall, 2009).

The contribution of my publications to equity evaluation is shown in Table 3.6. The main lesson from *Building an evidence base for community health* is that these services and practitioners mainly understand equity in terms of access to services for disadvantaged or marginalised groups of people. To evaluate equity issues more comprehensively, the evaluation needs to assess the impact of the program on different groups or populations to ensure the program is not contributing to inequity of outcomes as well as access (Kelly, et al., 2007). This is much more problematic than assessing access alone since community-based programs are seldom in a position to assess longer term outcomes and link these to the program. However,

program theories could be developed that encompass equity concepts and lay out the areas where potential equity factors need to be considered.

Table 3.6 Evaluation of equity concerns

Health promotion program	**Evaluation findings**	**How evaluated**
Eat Well SA	Programs directed to rural and remote communities; grants for low income and diverse cultural groups	Proportion of funding to disadvantaged communities Grant application process and feedback
What makes for sustainable Healthy Cities initiatives?	Not evaluated in this study	
Building an evidence base for community health	Reports focussed on access to services	Review of written program evaluation reports
Healthy Ageing – Nutrition	Program rationale suggests that older people are at nutritional risk	Not systematically evaluated

3.7.2 Practical and resource issues

The issues described above are exacerbated by lack of investment in methodological development, training and resources for health promotion evaluation. Health promotion programs, such as those that are the subject of my evaluations, are seldom generously funded and struggle to provide sufficient resources for evaluation (Baum, 2003; Potvin & Bisset, 2009). With regard to larger scale evaluation research, a review of Australian public health research funding (Public Health Research Advisory Committee, 2008) reported a lack of a strategic approach to funding of

public health research and serious concerns about public health research being disadvantaged (compared to medical and clinical research) in the grant application process. Further, an increased investment in public health research by the National Health and Medical Research Council was recommended. The complexity of community-based health promotion, as illustrated in my publications, suggests research and evaluation will need an increase in investment in order to further knowledge about how programs work and the evidence for effective practice.

While theory-based evaluation calls for input from all stakeholders, practitioners' priorities are centred on the demands of delivering services and programs and their capacity to undertake or be partners in evaluation is limited. The experience captured in my publications suggests that collaborative approaches to evaluation are likely to be more time and resource intensive for evaluators, and program stakeholders, than approaches with fewer stakeholders participating. For example, in *Healthy Ageing – Nutrition* the organisations demonstrated limited capacity to contribute to the broader evaluation. Health promotion initiatives such as Healthy Cities and specifically the *Eat Well SA* and *Healthy Ageing – Nutrition* projects aim to create change within organisations by increasing capacity for health promotion. The intention is to put health promotion on the agenda of an organisation that may not see this as its core role but which nevertheless does have an influence on the health of clients or the wider population. The extent to which the health promotion initiative has infiltrated an organisation's culture is a key focus for evaluation in this context. Indicators of embedment might include the level of commitment from organisational leaders, incorporation into strategic planning, the existence of new policies or procedures and resources to support implementation and staff development (Johnson & Baum, 2001). The *Eat Well SA* and *Healthy Ageing – Nutrition* evaluations investigated

some of these aspects. Longer term assessment of the sustainably of these changes was not able to be undertaken due to the fixed term of the evaluation.

Some theories underpinning health promotion, such as organisational change theory and development of healthy public policy, suggest that much health promotion activity (or advocacy for policy change) needs to be long-term before achievements are likely. Sustained change is more likely if the initiative involves structural, environmental, organisational and policy changes rather than only change in individual attitudes, knowledge, skills and behaviour (Labonte, 1992; Wilkinson, 1996). Using Hawe and colleague's (1997) three dimensions for capacity building for health promotion, Table 3.7 shows how these were constructed in projects reported on in two of my publications.

Table 3.7 Dimensions of capacity building

Capacity building dimension	***Eat Well SA***	***Healthy Ageing – Nutrition***
1. Infrastructure and service development	Increased skills and knowledge of importance of healthy eating; health promotion goals introduced; human, information and financial resources provided	Service and workforce development and support provided
2. Problem solving	Increased community and organisational skills for collaborative action, readiness to take collaborative action	Increased recognition of links between healthy ageing and nutrition
3. Sustainability	Organisational commitment; strategies owned and institutionalised	Policies and strategies embedded

Infrastructure and service development was able to be evaluated through measurement of the resources provided to participating organisations and groups, and increased awareness of the health promotion issues. Information resources comprised a large part of both projects and satisfaction and use was assessed, suggesting organisations valued and used these resources. Human resources were found to be important to both projects to provide drivers of change and ongoing support. This was particularly so in *Healthy Ageing – Nutrition* where no financial support was given to partner organisations. Overall, service development in the form of provision of resources enabled organisations to take on a new, or increased, health promotion role outside of their usual remit.

Having raised awareness of the issues, the next step in capacity building is for organisations to recognise that they can take action towards problem solving. For *Eat Well SA* much of this action was based on collaborative strategies that brought together stakeholders from different sectors or interests to share skills and perspectives. So, for example, nutrition professionals were able to work with communities and agencies to improve food supply to rural and remote areas. For *Healthy Ageing – Nutrition*, evaluation suggested that partner organisations recognised the role they could play in improving health in the aged population and acted to bring about changes.

3.8 Role of the evaluator

The final issue examined in this chapter is the changing role of the evaluator. As different approaches and methodologies have come to be used in evaluation research, so the role of the evaluator has changed over time. The evaluator role reflects the definition or understanding of the purpose of evaluation. My role as illustrated in my

publications has also varied over time and in response to the demands of the consultancy context. Drawing on a range of evaluation literature in Chapter 2 and my own experience, I summarise eight different evaluator roles along a continuum (see Table 3.8). These roles are described more fully below.

Table 3.8 Evaluator roles

Descriptor	**Role**
Technical measurer	measures the extent of achievement of pre-determined outcomes
Describer	describes the initiative and the apparent outcomes
Judge	judges worth or against standards
Hypothesis tester	tests the hypotheses upon which the program is based
Negotiator	creates a consensus of constructed reality
Theory provider	fills the gaps in the theory of action
Mediator	creates consensus among interest groups
Partner	embedded within the program

The technical measurer is seen as a value and content free role, the only requirement of the evaluator is competence in measuring the extent of achievement of pre-determined outcomes. The describer role is also seen as value-free, here the evaluator describes the initiative and the apparent outcomes arising.

In the role of judge, the evaluator may be a judge of worth where 'evaluation entails making informed judgements about a program's worth, ultimately to promote social change for the betterment of society' (Grembowski, 2001 p13) or a judge against a standard or standards where 'Intrinsic to evaluation is a set of standards that (explicitly or implicitly) define what a good program or policy looks like and what it

accomplishes' (Weiss, 1998 p320). This judgement role assumes an objective, value-free evaluator who bases their judgement on specialist knowledge or agreed standards. For the hypothesis tester, the role of the evaluator is to test the hypotheses upon which the program is based (Green & Kreuter, 1999). An agreed set of testable hypotheses and an objective evaluator are assumed.

Guba and Lincoln (1989) describe the evaluator as a negotiator, and the role is to create a consensus of reality and values among all the stakeholders. In this approach, evaluation is 'a process whereby evaluators and stakeholders jointly and collaboratively create (or move towards) a consensual valuing construction of some evaluand' (Guba & Lincoln, 1989 p263). In this role, the values and characteristics of the evaluator become more explicit and may or may not carry greater weight than that of other stakeholders. In the theory provider role the evaluator fills the gaps in the theory of action, that is describing what must happen to get to the next stage of the program (Patton, 2002). In this role the evaluator draws on their own values and expertise, but may also enlist other stakeholders, and recognises the diversity of values and interests. For the evaluator as mediator, the role is as source of information to stakeholders, a negotiator and consensus generator, bringing professional expertise to mediate between different stakeholder interests (Chen, 2005). In this role the evaluator is an active member of the program team, addressing issues of power differentials and relative stakes.

Finally, the evaluator may act as a full partner. In this role the evaluator is embedded within the program and is a partner to the evaluation user or program personnel (Patton, 2011). This role requires the evaluator to bring evaluation thinking to the program stakeholder group while supporting the values and vision of the program. Of

course, for a particular evaluation, any of the above roles may be appropriate. Further, these roles are not mutually exclusive, and may overlap, so that a mix of roles is likely.

My evaluation work illustrates different roles on the continuum and the mix of roles that I have held. My role has tended towards the latter end of the continuum as my experience and confidence as an evaluator has developed, and congruent with developments in evaluation theory and practice. So, for example, in the *Eat Well SA* evaluation, my role was part describer of what took place and the outcomes achieved, part judge of worth and part tester of the hypotheses on which the program was based. The evaluation was conducted according to the contractual agreement with the funders and with only a little room for innovation or adaptation. For the *What makes for sustainable Healthy Cities initiatives?* paper my role as a theory provider was to articulate the implicit theory about sustainability. The long association of all the authors of this study with Healthy Cities Noarlunga meant there was a shared value-base and understanding about the project on which to build. For *Healthy Ageing – Nutrition* the role was more of a mediator, outside the project team but closely aligned. Thus, the evaluation was able to draw on data from the project manager's reflexive practice and also gave the partner organisations an opportunity to make their perspectives heard outside of the project structure.

As my evaluation work describes above, moving along the continuum changes the evaluator role from that of a content and value-free measurement approach to a more collaborative exercise to identify and build consensus about underlying program theory and to work in partnership with the program stakeholders. As we move along the continuum, the evaluator needs more content knowledge and understanding of

the program and stakeholders' values and perspectives. This is illustrated by the Healthy Cities and *Healthy Ageing – Nutrition* evaluations where there was close engagement with the programs. Chapter Two concluded that community-based health promotion is underpinned by multiple theories and models but with some common principles of practice. This means that the evaluator's role should include facilitation of theory consensus, an awareness of health promotion theory and practice, and should support and reflect health promotion principles. In evaluation of community-based health promotion initiatives the complexity of the setting, with many stakeholders and power differentials, suggests that the negotiator/mediator role, working in partnership, is likely to be most effective in uncovering multiple perspectives and in reflecting health promotion principles of empowerment and participation.

The role played will be both constrained and mediated by the circumstances of the evaluator's engagement and organisational position. Evaluators can be engaged: as an internal evaluator (personnel and resources from the organisation or project); as a contracted consultant (usually for profit); from a government agency; through an academic /research consultancy (may be for profit or funded through a grant). My position in SACHRU has been as an academic researcher within an organisational unit of a university. SACHRU has to account for the part-funding it receives from the State Government Health Department and mostly adheres to a commercial costing scheme for consultancy services provided from within the university. Thus my evaluation role is usually undertaken as a consultant acting within an agreed contract and budget from the program management, for example *Eat Well SA* and *Healthy Ageing – Nutrition*. I have found that evaluation contractors and managers of health promotion programs are usually sympathetic to the concept of evaluation grounded

in the health promotion principles and approaches I have described. At the same time, however, contracts are bounded by time and resource constraints and the practical outcome of this is that priorities have to be set and decisions made according to what is realistic. Thus, evaluation design is a balance between the ideal and the pragmatic with the focus generally decided according to the purpose of the evaluation.

The research leading to my other three publications has been resourced through core funding and therefore forms part of SACHRU's service agreement i.e. *What makes for sustainable Healthy Cities initiatives?, Building an evidence base for community health* and *Framework and tools.* In this context the research design is less constrained by a specific program and its resources. This increases opportunities for theory development but also removes the evaluator from direct involvement with program practice so care needs to be taken to ensure relevance to the health promotion practice community. For these studies, my co-researchers and I have placed emphasis on engaging with stakeholders to increase the relevance and usefulness of findings.

3.9 Summary and conclusions

We have seen that health promotion activities tend to be poorly funded compared to care and treatment and this difference extends to resources for research and evaluation. A further resource issue is that health promotion programs and their evaluations suffer from short-term funding (as noted with *Eat Well SA* and *Healthy Ageing – Nutrition*). The capacity of health promotion practitioners and services to undertake or contribute to evaluation is limited and there is a need for evaluation support and resources. These could be provided from inside the organisation or

through a consultancy service such as SACHRU. Since there is limited value in one-off evaluations, particularly of small scale programs, an institutional support system would enable consolidation of findings and a cumulative approach to knowledge about health promotion evaluation. This would assist with the ongoing challenge for health promotion evaluation design – the tension between the need for generalisable knowledge about good practice and the context specific setting for most programs.

In terms of evaluation design, choice of appropriate short and long-term outcomes and methods to measure changes is critical. Long term changes in health status are unlikely to be measurable in a specific recipient population of a health promotion program, as opposed to the general population and, even if change is seen, causal links are problematic. A more realistic approach is to focus on shorter term outcomes such as knowledge, skills and behaviour change and draw on theory or evidence from other programs about the likelihood of these contributing to improved health. The links between these short and longer term outcomes can usefully be expressed in a program logic model. Particularly for more complex interventions, the program theory and logic model is likely to be emergent and subject to change as the context and players interact and change.

Since community-based health promotion programs operate in an open and dynamic environment, it is also important to document change in the social, economic and political context and assess the influence of this on the way an initiative unfolds and its outcomes. This is a gap in my published evaluations, although there is some focus beyond individuals to organisational and policy change, capacity building of individuals and organisations, and sustainability of effective programs.

This chapter has illustrated the development of my evaluation role through analysis

of the five publications constituting part of this thesis. In that time, my evaluation approach has moved from describer and judge to using stakeholder perspectives to articulate a program theory and to increase understanding of the importance of context. I have taken the position that, for evaluation of community-based health promotion initiatives, the principles of community participation, equity and partnership approaches should be enshrined as part of the evaluation design and I have developed and strengthened these aspects of my evaluation practice. At the same time, I have needed to juggle the ideal with the practical constraints of time, resources and contractual accountability.

I have taken the standpoint that empowerment of individuals and communities is an underpinning principle of health promotion theory and practice. This means that evaluation of health promotion programs should also be empowering for the program recipients and other stakeholders. A number of implications from this are apparent: the health promotion program and its evaluation design need to be flexible in order to allow for, and respond to, stakeholder input; the evaluator needs to take on a negotiator or mediator role in order to bring different interests to a workable consensus and stakeholders need the skills, resources and interest to be engaged meaningfully in the evaluation. It is also clear that the evaluation should be considered at the planning stage of a program so that the evaluator can bring their skills to assist stakeholders to conceptualise and articulate the program theory. These factors together should help to increase the level of control held by program recipients and other stakeholders and the relevance and usefulness of the evaluation findings.

The literature review in Chapter Two on the development and evaluation of

community-based health promotion initiatives and my publications demonstrate some of the complexities apparent in both the initiatives and the environments in which they are implemented. These include: lack of agreed theory and program logic models; context-contingent nature of programs; frequent change and turbulence in community context and players; multiple stakeholders, goals and strategies; and uncertainty of outcomes. The extent of complexity characteristics for my three published program evaluations is outlined in Table 3.9.

Table 3.9 Characteristics of complexity for three program evaluations

Program evaluation	Uncertainty of outcomes	Non-linear	Dynamic	Adaptive
Eat Well SA	moderate	moderate	high	low
What makes for sustainable Healthy Cities initiatives?	high	high	high	high
Healthy Ageing – Nutrition	high	moderate	high	high

'Uncertainty of outcomes' refers to the degree of unpredictability of processes and outcomes from the program. The greater uncertainty for *What makes for sustainable Healthy Cities initiatives?* and *Healthy Ageing – Nutrition* is reflective of the participatory planning in these programs.

'Non-linear' refers to the extent to which the program logic and implementation of the program is affected by feedback from the context. Healthy Cities initiatives, with a focus on intersectoral activity and actions across diverse environments, require flexible logic modelling in order to reflect the changing context for the program.

‘Dynamic’ programs are those with have a high degree of responsiveness to emerging interactions and relationships. All three programs relied heavily on building and maintaining collaborative partnerships, often with new players where interactions are particularly unpredictable.

‘Adaptive’ refers to the extent to which the program is able to adapt to changes both within the implementation space and in the wider environment. *Eat Well SA* was less adaptive as it was planned with a set of objectives and strategies from a bureaucratic base and had limited capacity to evolve during the timeframe of implementation. Both *What makes for sustainable Healthy Cities initiatives?* and *Healthy Ageing – Nutrition* had a more flexible base, while the sustained timeframe for Healthy Cities contributed to a high capacity, and need, for adaptation.

The next chapter considers complexity theory as a way to assist in thinking about the evaluation of programs with the characteristics outlined above. Complexity theory and the arising notion of developmental evaluation may help to address some of the challenges that have been described for evaluation of community-based health promotion initiatives.

CHAPTER FOUR: THE EMERGING AREA OF COMPLEXITY THEORY AND DEVELOPMENTAL EVALUATION

4.1 Introduction

So far, this thesis has considered the context leading up to my publications and how this has influenced my evaluation work. Chapters Two and Three identified the challenges and issues arising from the context in which community-based health promotion evaluation sits. These issues include the inappropriateness of experimental methods in a dynamic community context, consequent problems of causation, attribution and transferability of findings, diversity of population groups and implementation strategies and the importance of the evaluation reflecting health promotion principles of equity, working in partnerships and community participation and empowerment. This chapter specifically addresses Research Question Four - What are the overall lessons from the evaluation practice presented in my publications and how do they inform new approaches to evaluation of community-based health promotion initiatives? - by drawing together the lessons from my published evaluation work and examining recent approaches to evaluation in order to assess their potential contribution to deal with some of these issues.

As previously discussed, community-based health promotion initiatives are often described as complex or as having complex components. Table 3.9 presents a complexity classification of the projects evaluated in my publications. The emergence of complexity theory and related literature since this evaluation work was undertaken appears to hold promise in meeting some of the evaluation challenges I have identified. I therefore review this developing field to explore its potential to

enhance community-based health promotion evaluation and test its application on my evaluation work in particular.

Firstly, categorisation of simple, complicated and complex interventions is discussed. Complexity theory is then outlined in order to set the context, and its synergy with health promotion explored. I consider the potential use of complexity theory in evaluation of community-based health promotion initiatives and draw on one of my publications to assess the contribution that complexity theory might bring to health promotion evaluation. A recent approach – developmental evaluation – is analysed as an example of how complexity theory has been applied to the field of evaluation. Finally, I draw on the literature, my published works and the lessons from these to construct a conceptual model that illustrates my thinking in combining complexity and program theory and how this might assist the process of evaluating complex initiatives in complex settings.

4.2 Complex social problems and interventions

Social problems and social interventions (such as community-based health promotion initiatives) can be considered as inhabiting simple, complicated, complex or chaotic spaces (Patton, 2011). A number of commentators have developed frameworks and models that can help us to understand complexity by describing these spaces and the types of knowledge that pertain to each.

Plsek and Greenhalgh (2001) present a framework for thinking about simple, complex and chaotic knowledge approaches according to the level of certainty about what the problem is and the level of agreement on what to do about it (Figure 4.1). In this framework, simple knowledge is appropriate when there is high agreement about the nature of the problem and what to do about it. As certainty and agreement lessen,

we encounter the area of complexity. At the extreme, there is such low agreement about the problem and what to do about it that we enter the area of chaos.

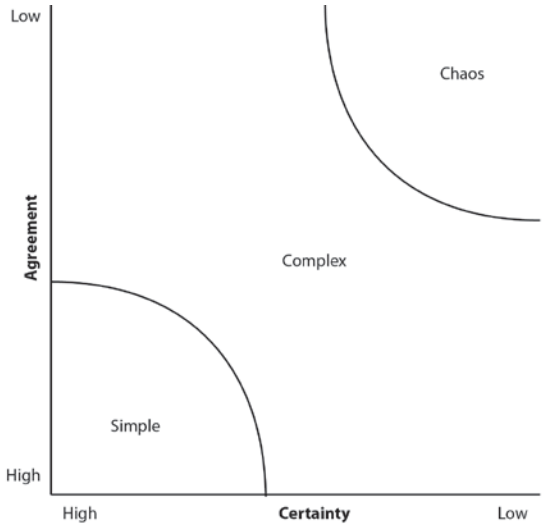

Figure 4.1 Simple, complex and chaotic knowledge framework (adapted from Plsek & Greenhalgh, 2001 p627)

Patton (2011) introduces an area of complicated problems and further distinguishes between technical and social complication: technically complicated problems have low certainty but high agreement, whereas socially complicated problems have high certainly but low agreement (see Figure 4.2.) Technical complexity requires coordination and high levels of expertise; social complexity involves multiple stakeholders with different perspectives and values. This description suggests that many health promotion interventions will fall into the socially complicated area. There may be cross-sector partnerships and community engagement bringing multiple stakeholders perspectives. While there may be agreement about long-term outcomes, there is likely to be a range of views about how to achieve these. Aspects of the Healthy Cities and *Eat Well SA* evaluations demonstrate this type of

complication.

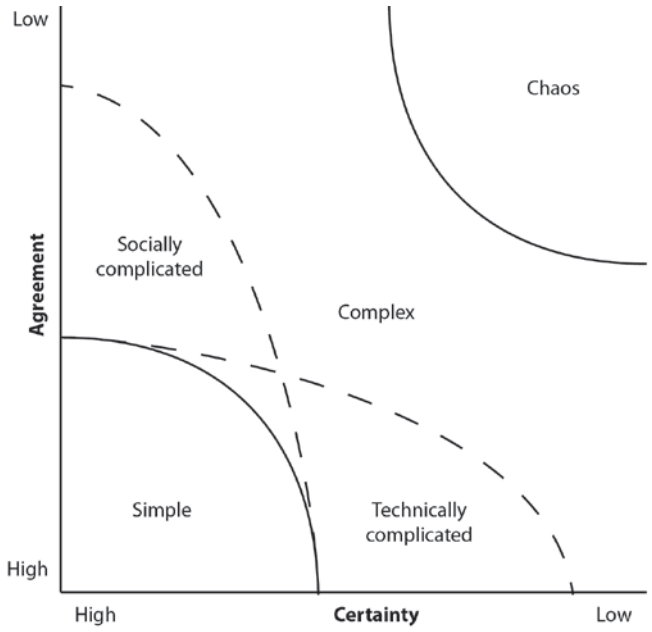

Figure 4.2 Simple, complicated, complex and chaotic zones (adapted from Patton, 2011 p 94)

A useful example of simple, complicated and complex activities is given by Westley and colleagues (Westley, Zimmerman, & Patton, 2006), see Table 4.1. Their description of complex problems resonates with much community-based health promotion: implementation is not standardised, successful outcomes do not necessarily translate to a new setting, responsiveness to the community and the setting is needed as each is different, there is uncertainty of outcomes and a holistic approach is required that goes beyond discrete program components.

Table 4.1 Simple, complicated and complex problems (Westley, et al., 2006 p9).

Simple: baking a cake	**Complicated: sending a rocket to the moon**	**Complex: raising a child**
The recipe is essential	Rigid protocols are needed	Rigid protocols have limited application or are counter-productive
Recipes are tested to ensure easy replication	Sending one rocket increases likelihood of success with the next	Raising one child provides experience but is no guarantee of success with the next
No particular expertise required, experience increases success rate	High levels of expertise and training in a variety of fields are necessary for success	Expertise helps but only when balanced with responsiveness to the particular child
A good recipe produces nearly the same cake every time	Key elements of each rocket much be identical	Every child is unique and must be understood as an individual
The best recipes give good results every time	There is a high degree of certainty of outcome	Uncertainty of outcome remains
A good recipe notes quantity and nature of parts needed and specifies order in which to combine them, but room for experimentation	Success depends on a blueprint that directs development of separate parts and specifies the exact relationship in which to assemble them	Can't separate the parts from the whole; essence is the relationship between different people, experiences and moments in time

Thus, simple interventions (or simple components of interventions), for example, immunisation, generally have a single agreed set of objectives and anticipated outcomes, are governed by a single organisational unit, have a standardised delivery or implementation, are necessary and sufficient to produce the expected outcomes

and work in the same way for everybody, and have a simple cause-effect relationship (Funnell & Rogers, 2011). Complicated interventions usually involve a number of interrelated parts, all of which need to function together in a predictable way (Ling, 2012). They may have different objectives valued by different stakeholders, multiple, competing objectives and objectives at different levels in the system. Interventions may be governed in a partnership with several organisations all with their own requirements and values and are likely to be adapted for different groups as they are implemented (Funnell & Rogers, 2011). The intervention may be only one of several ways to achieve the desired impact, it may work in conjunction with other interventions, work only for specific people or in favourable environments. The relationship between cause and effect requires expertise to understand and predict (Funnell & Rogers, 2011). The social marketing campaign component of *Eat Well SA* could be seen as complicated since different populations may react in different ways, the campaign may form part of an overarching strategy and cause and effect are not readily apparent.

Complex interventions are characterised by feedback loops, adaptations and multiple components (Ling, 2012) that may act independently or interdependently (Campbell et al., 2007). They have emergent objectives in response to changing needs, opportunities and challenges and their governance involves an emerging cast of partners and relationships. The changing nature of the adaptive intervention means that cause and effect relationships cannot be predicted, but they may be interpreted retrospectively (Funnell & Rogers, 2011). Using these characteristics it is clear that many community-based health promotion initiatives, for example Healthy Cities, are complex since their development is responsive and adaptive in relation to stakeholder and environmental factors changing over time. These complexity factors mean that

static, linear approaches to evaluation will miss much of the adaptive nature of the intervention and may focus on issues that are no longer relevant to many stakeholders.

Classification into simple, complicated and complex is, however, not straightforward. The term ‘complex’ is used broadly in everyday language to describe anything that goes beyond very simple characteristics. In the health field an example of a rather loose use of terminology is the Medical Research Council guidelines (Medical Research Council, 2008). Here the description of a complex intervention is more like a complicated intervention (Ling, 2012; Shiell, Hawe, & Gold, 2008) and this has implications for the way evaluation is conceived. Thus non-experimental methods are described as weak and it is stated that before evaluation the intervention must be developed to the point where it can reasonably be expected to have worthwhile effect. Evidence from recent high quality systematic reviews is recommended. The guidelines acknowledge there may be a need for adaptation to circumstances but are concerned about the loss of program fidelity this brings. This approach does not seem to leave much room for the creative, adaptive or emergent design that complexity theory would suggest. Again, the evaluation may lose relevance and utility if it is focussed on outcomes from delivery of a standardised program when in fact the program is developmental.

Distinguishing between simple, complicated and complex interventions can provide a framework for developing the program theory and deciding what factors of the evaluation should be prioritised and what evaluation questions to ask (Funnell & Rogers, 2011; Patton, 2011). So, for simple interventions, a linear logic model is appropriate; for complicated interventions, system diagrams or maps that show

relationships are suggested; and for complex interventions, developmental evaluation that is flexible, emergent and dynamic is better suited to complex interventions (Patton 2011).

Notions of complexity are then a potentially useful tool for determining the type of evaluation required for community-based health promotion initiatives. The next section considers what the use of complexity theory might bring to health promotion evaluation.

4.3 Complexity theory and health promotion evaluation

Complexity theory is a recent addition to the field of evaluation. To date, it has been mainly used in organisational strategy, knowledge management and cultural change (Kurtz and Snowden 2003) however the concepts are relevant to health promotion evaluation. This section discusses the synergy between complexity theory and health promotion and presents complexity as a potentially useful new approach to evaluation of community-based health promotion initiatives. Complexity theory provides new insights into the challenges identified in my published work by recognising the critical role of relationships and interactions in the community setting and the unordered nature of these aspects of a health promotion program.

4.3.1 Complexity theory

While there is little agreement about a definitive complexity theory (Nunn, 2007; Patton, 2011) there is consensus on a number of elements that go to make up complex systems. In general a complex system is:

- Non-linear
- Emergent/self-organising

- Adaptive
- Unpredictable/uncertain
- Dynamic
- Co-evolutionary/ dependent on history (Nunn, 2007; Patton, 2011; Shiell, et al., 2008)

Complexity theory helps to illuminate characteristics of problems and interventions that fall into the complex zone and is grounded in the assumption that not everything is knowable or predictable in the scientific sense. Complexity theory holds that human society in particular is an un-ordered (or emergent), dynamic system in which it is not possible to predict, with any certainty, behaviour change or its impact.

The Cynefin framework (Kurtz and Snowden 2003) (see Figure 4.3) was originally designed for use as a group exercise for collective decision making about complex systems. This framework is a useful way to categorise cause and effect patterns relating to human behaviour change. There are four domains, two ordered and two un-ordered. In the ordered world, patterns of cause and effect are known, or knowable given sufficient resources. In the un-ordered world, patterns are complex and arise from dynamic interactions. Cause and effect cannot be predicted and, indeed, the act of investigation may lead to new interactions and new patterns. Feedback loops can positively drive change or negatively stifle change (Alvaro et al., 2010). Chaos describes the quadrant where there is such a high level of turbulence that cause and effect cannot be perceived.

The Cynefin framework resembles the certainly/agreement matrix with ‘known’ corresponding to simple and ‘knowable’ corresponding to complicated. The areas of transition between quadrants represent critical points for change, with the interface at

the ‘edge of chaos’ the space where there is greatest possibility of substantive change and innovation (Alvaro, et al., 2010; Martin & Félix-Bortolotti, 2010).

Un-ordered – **complex**, patterns emerge through interaction, cause and effect perceived but not predicted and have retrospective coherence. We cannot be sure patterns will repeat because sources of patterns are not open to inspection (and this may itself disrupt patterns)	**Ordered** – **knowable**, cause and affect separated over time and space but could be known with sufficient time and resources, instead rely on expert opinion. Reductionist
Un-ordered – **chaos**, turbulence, no cause and effect perceivable	**Ordered** – **known**, linear cause and effect, not disputed, repeatability, predictive models. Best practice can be described

Figure 4.3 The Cynefin framework (based on Kurtz & Snowden, 2003)

4.3.2 Complexity theory and implications for health promotion evaluation

The concept of human interaction taking place in a dynamic system has implications for health promotion evaluation. Health promotion has been described as ‘systems thinking in action’ (Norman, 2009 p869) and Norman argues that, as the field of health promotion has developed, a more complex picture of health and health promotion has emerged. Tremblay and Richard (2011) see a convergence of ideas in complexity theory and health promotion as the complexity of social structures and change is being recognised. The synergies between complexity theory and health promotion include a concern for an integrated holistic approach and the need for a comprehensive, dynamic, non-linear understanding of issues such as legitimising lay

knowledge and participation (Tremblay & Richard, 2011).

Complexity theory challenges the notion that linear cause and effect relationships can always be found. Kurtz and Snowden (2003) question the assumption that human decision making and policy making is based on order, rational choice and intentionality. Kurtz and Snowden (2003) argue that these assumptions are based on enlightenment science that distinguishes order from chaos and maintains a belief that all things can be (eventually) known. However, in some situations 'the lack of order is not a matter of poor investigation, inadequate resources or lack of understanding, but is a priori the case – and not necessarily a bad thing, either' (Kurtz & Snowden, 2003 p464). This is pertinent to evaluation of community-based health promotion where human behaviour sets the context and a linear relationship between cause and effect, rational choice and intentionality cannot be assumed. Hawe, Shiell and Riley (2009) describe interventions as events in a system that either 'leave a long-lasting footprint or wash out depending on how well the dynamic properties of the system are harnessed' (p 270). A health promotion 'complex intervention' has interacting components similar to those described by complexity theory: discretionary behaviour or actions; multiple levels of individuals and groups; and the need to be flexible (Hawe, Shiell, et al., 2009). Most significant is the setting context, interaction between setting and intervention, and relationships (Hawe, Shiell, et al., 2009; Matheson, Dew, & Cumming, 2009). The *Eat Well SA* project, for example, worked across multiple levels with interacting individuals and groups in settings such as child care, schools, families and community organisations. Thus, a systems approach starts with context, and the intervention is a way to create new roles and increase interaction between players. This is similar to the context-mechanism-outcome configuration in the realist approach of Pawson & Tilley (1997). Westhorp (2012)

suggests that complexity theory adds value to realist evaluation by making clear what factors to look for in the initial context and what factors act as controlling parameters in tackling resistance to change. This is similar to Signal and colleagues' (2012) argument that complexity can help identify the best points for intervention within self-organising, stable systems. For example, a food security system map showed key areas for intervention as money available in households and the cost of food (Signal, et al., 2012).

Kurtz and Snowden (2003) further argue that ordered systems thinking is reductionist:

> 'Ordered-systems thinking assumes that through the study of physical condition, we can derive or discover general rules or hypotheses that can be empirically verified and that create a body of reliable knowledge, which can then be developed and expanded.' (Kurtz & Snowden, 2003 p466)

This reductionist approach attempts to optimise the system by optimising the parts within it. However, in dealing with un-order domains, the whole is not simply the sum of the parts because any act may change the nature of the whole system (Kurtz & Snowden, 2003). For example, the complexity of health promotion is illustrated by the interconnectedness of issues that impact on health and wellbeing (Norman, 2009). Thus, a state of health or ill health is rarely dependent on a single cause, but rather, on an interplay of biological, behavioural, social and environmental factors. Norman argues that health promotion is not centrally controlled or delivered through hierarchical structures, rather it works through self-organisation and social networks with unclear boundaries, diverse actors and non-linear change (Norman, 2009). This description particularly fits with community-based health promotion initiatives. Understanding that communities and settings are 'un-ordered' leads to the notion that interventions may act in unanticipated ways. A holistic approach to evaluation is

required that monitors all changes in order to capture unanticipated outcomes, rather than focussing solely on looking for predicted change. Complexity theory presents an opportunity to adopt a dynamic ecological approach to research and evaluation in community based interventions (Hawe, Shiell, et al., 2009). In the *Eat Well SA* evaluation for example, ideas from complexity would have led to much closer scrutiny of the food environment, how this changed over the course of the project and the impact of this on what the project was able to achieve.

The Cynefin framework presented above has been used by Patton (2012) to delineate the evaluation role. Figure 4.4 is based on Patton's work with health promotion examples added. Thus, the role of evaluation in the simple domain is to monitor achievement against known best practice. In the complicated domain, evaluation draws on expert knowledge to predict effective practice and look at the impact of context on cause and effect patterns. In evaluation of complex interventions, the context is monitored to see what emerges from dynamic interaction. Feedback and reflection are crucial to the process. In the simple domain evaluation aims for certainty, in the complicated domain to reduce known uncertainty and in the complex domain to support self-improving systems (Ling, 2012).

Un-ordered – complex. Cause and effect perceived but not predicted. Patterns emerge through interaction. High disagreement about problem and what to do about it. High uncertainty about how to produce desired outcome. **Evaluation**: identify initial conditions and monitor what emerges; provide ongoing rapid feedback; track actions and decisions; facilitate reflexive practice; embed evaluation in intervention process. **Health promotion example**: *health inequity intervention such as Healthy Cities*	**Ordered – knowable**, **complicated.** Cause-effect is context-contingent Some disagreement about problem and what to do about it. Uncertainty about attaining desired outcome. **Evaluation**: validate effective practice with attention to context; provide testable theory of change from expert opinion; report cause-effect complications and implications. **Health promotion example**: *healthy eating education programs*
Un-ordered – chaos. Turbulence, no cause and effect perceivable. Pattern detection unreliable. High conflict and uncertainty. Dynamic interactions hard to follow. **Evaluation**: distinguish between better and worse data; interpret with caution; find activity where evaluation can make an immediate contribution to surviving chaos. **Health promotion example**: *interventions to mitigate health impacts of climate change*	**Ordered – known**, **simple.** Clear linear cause and effect. High agreement about problem and what to do about it. Cause – effect link predictable and controllable. **Evaluation:** validate best practice; monitor implementation of best practice; report deviations from best practice and implications. **Health promotion example:** *fluoridation of drinking water*

Figure 4.4 The Cynefin framework and health promotion evaluation (based on Patton 2011 p 109) with health promotion examples added

4.4 Using complexity theory to address evaluation challenges

In the section above I have argued that complexity theory provides a means of addressing some of the evaluation challenges associated with community-based health promotion efforts and illustrated in the discussion of my publications. This section describes the characteristics of complexity and the implications for evaluation and Table 4.2 builds on and extends Patton's ideas to link complexity characteristics with health promotion and details the implications for health promotion evaluation.

Table 4.2 Complex health promotion initiatives and evaluation (adapted from Patton, 2011 p150-151)

Complexity characteristic	**Health promotion complexity**	**Implications for health promotion evaluation**
Non-linearity: Sensitive to initial conditions, small actions can stimulate large reaction, tipping point	Initiatives act in the community. Initial and changing context effects how initiative takes place and how it influences participants	Linear program logic models & plans do not necessarily reflect what actually happens. Evaluation needs to record initial context and monitor changes
Emergence: Patterns emerge from self-organisation among interacting agents. Each agent has own path but interacts with others' paths. Interactions cohere, becoming greater that the separate parts.	Interactions between stakeholders and sub-systems of the initiative lead to new ways of implementation. Population sub-groups may experience and respond to the initiative in different ways. Outcomes will not always be apparent at start of initiative and may change developmentally.	Track emerging interactions and networks between stakeholders, differing experience and outcomes for different sub-groups (i.e. equity concerns). Look for unanticipated events. Evaluation design should be emergent in parallel with the initiative. Evaluation of whole is more than sum of components.

Adaptive: Interacting elements and agents respond and adapt to each other and to environments. What emerges is a function of ongoing adaptation.	Uncertainty and unpredictability in how the initiative will take place means initiatives adapt to stakeholders' experiences and changes in the context.	Capture perspectives from different stakeholders, feedback evaluative information to all groups. Evaluation design should be adaptive
Uncertainty: Processes and outcomes are unpredictable, uncontrollable, and unknowable in advance	Community participation and changing contexts mean that planned processes and planned for outcomes will very likely be subject to revision.	Identify sources of uncertainty, disagreements and turbulence. Resist forcing order and control, imposing linear logic models and predetermined outcomes. Anticipate unexpected events and provide rapid feedback.
Dynamical: Interactions within and between parts of systems can be volatile, changing rapidly and unpredictable due to interdependence of key factors and variables.	Health promotion initiatives may be subject to changes in key personnel, political focus, new policy and intra and intersectoral events.	Track how and why changes in interactions between stakeholders and sub systems occur. The evaluation should be prepared to capture volatile and turbulent change.
Co-evolutionary: As interacting and adaptive agents self-organise, ongoing connections emerge and agents co-evolve as parts of the system over time.	Sustainable health promotion initiatives arise from ongoing connections and system development such as policy change, legislation and increased community capacity for action.	Evaluation not independent but co-created with the initiative, through feedback and facilitation. Process evaluation affects initiative development. Include in evaluation design, participatory and consultative process about how initiative will be rolled out.

Each of the complexity characteristics in Table 4.2 demonstrates lessons for community-based health promotion evaluation.

Non-linearity

In dealing with the non-linearity of community-based health promotion initiatives and longer time frames of much health promotion activity, complexity theory supports a non-linear approach with emphasis on observation, analysis and understanding of the pre-intervention context. Indeed Hawe, Shiell and Riley (2009) contend that it is the setting that is complex rather than the intervention and this requires a move from program evaluation to context evaluation. I argue in this thesis that both the intervention and the context are complex and therefore both types of evaluation are needed.

Emergence

Traditional evaluation methods often test each component of an initiative in isolation when complexity theory tells us that these are interacting and contribute to the whole through emerging patterns. According to complexity theory, therefore, evaluation needs to be holistic and to consider how component parts interact to contribute to the observed changes. As Ling describes, complexity theory leads to thinking of an intervention as 'including a process of reflection and adaption as the characteristics of the complex systems become more apparent to practitioners' (Ling, 2012 p84). Evaluation then aims to understand the characteristics of the system and how effectively the program has been adapted. Showing how a particular system functions and how systems interact can assist with taking evaluation beyond the notion that each context is different (Ling, 2012). Ling gives an example of the many systems interacting in the prevention of deaths from cardio-vascular disease (disease factors, patient beliefs and behaviours, access to services, health system organisation,

public policy, socio-economic status) and notes that any intervention would need to interact with all of these. Thus a setting or context for health promotion is often a set of interacting systems and complexity thinking can show how each system functions and interacts by mutually supporting or undermining other systems affecting the intervention. According to Ling (2012) this increased understanding leads to the mid-range theories that can assist with generalisation of findings.

Adaptive

Complexity theory holds that agents in the evaluation respond to each other and to the environment in adaptive ways. This means it is important to capture stakeholder perspectives and understandings of the initiative. Drawing on stakeholders perspectives allows the evaluation to move from a focus on scientific knowledge to learning and increased capabilities for all participants (Hawe, Bond, & Butler, 2009).

Martin and Félix-Bortolotti (2010) identify two sub-systems to health care: simple/complicated, where efficiency and best practice are measured, and complex, where the focus in on functioning of relationships and positive and negative feedback loops. These authors position most health care evidence into the ‘simple’ domain, using randomised controlled trials to drive evidence-based medicine, however, health promotion and population health literature take a more complex, theoretical approach (Martin & Félix-Bortolotti, 2010). They suggest that health care is moving from the currently dominant reductionist paradigm to an edge of chaos transition. This is where innovation and creativity are likely to emerge.

Uncertainty

Uncertainty about processes and outcomes means that the program theory and program logic models are likely to be subject to change and stakeholders will need rapid feedback to deal with these changes. Complexity theory can assist by

recognising health promotion interventions as complex adaptive systems these can be strengthened and improved through reflection and feedback from formative application of complexity theory (Hawe, Bond, et al., 2009).

Dynamic

Unpredictable interactions within an initiative will lead to changes in networks, distribution of resources and power. So, for example, capacity building assessment should include not just measurement of the spread and uptake of change but should assess enablement, or improvement, of the structural position of people and organisations in the intervention network (Hawe, Shiell, et al., 2009) and how they relate to each other.

Co-evolutionary

A self-organising and adaptive system will have an impact on how an initiative or program is developed in the local context and this is a problem for traditional evaluation requiring intervention fidelity and standard forms of programs. Complexity theory suggests that this can be managed through minimum specifications rather than prescriptiveness and, certainly in community-based health promotion, rigid adherence to implementation is not helpful because the program needs to adapt to local context. To overcome this issue, Hawe and colleagues (2009) propose that the function of the intervention is standardised and fidelity is maintained to the theory rather than implementation.

4.5 Complexity theory and developmental evaluation

The relatively recent introduction of complexity thinking into evaluation means that there are few examples of its use in the field. This section considers a recent evaluation approach – developmental evaluation (Patton, 2011) – that draws on

complexity theory. Michael Quinn Patton is an eminent evaluation practitioner and theorist who has made many important contributions, the most recent of which is the idea of developmental evaluation. It is presented here to illustrate some practical applications of complexity theory and because it has contributed to my thinking about the evaluation model that I develop in the next section.

In his latest book, Patton (2011) details 'developmental evaluation' an approach that appears to show great promise for community-based health promotion evaluation as it embraces adaptation to a changing environment and identification of underpinning program principles that could be transferred to a new situation. It is therefore highly relevant to my work as it supports evaluation of programs that change in response to community input and provides a means of transferring findings to new contexts. Two of Patton's purposes for development evaluation are particularly pertinent to health promotion. Firstly, when a program is adapting in a complex situation, developmental evaluation can be used to identify principles to inform that ongoing development. Secondly, developmental evaluation can assist in adapting the general principles of a program to a new context. So, rather than adopting best practice without regard for local context, validated principles are adapted. Hawe, Shiell and Riley (2009) describe a similar notion of 'fidelity to theory' rather than fidelity to implementation.

Developmental evaluation is not intended to replace formative and summative evaluation but is proposed as a useful approach when an intervention or program has uncertain processes and outcomes that are unpredictable, uncontrollable and unknowable in advance; characteristics of many community based health promotion initiatives. Patton (2011) argues that developmental evaluation can adapt to the

realities of complex non-linear dynamics, rather than trying to impose order as traditional evaluation does. Thus, developmental evaluation is not a specific method but an approach to evaluation embedded with notions of complexity. Patton provides an example of how systems thinking can have an impact on how an intervention is implemented and evaluated. The intervention described aimed to address issues of obstructive sleep apnoea in police officers. A linear logic model for the program encompassed recruitment of police officers, an education component on what to do about sleep apnoea, and agreement to screening and treatment if necessary, all leading to healthier sleep patterns and higher level, safer performance. This assumes a rational decision making process by the police officers as they experience the program. However, in reality, decisions by the police officers are influenced by a network of relationships including work colleagues and supervisors, family members, personal doctor and so on. Developmental evaluation therefore considers both the linear model and the network factors as they influence the intervention's implementation and outcomes at individual, organisational and societal level. Similar scenarios existed in *Eat Well SA* and *Healthy Ageing – Nutrition*, with a network of relationships superimposed on the linear steps described in the program logic. While some of this network influence could be identified in the assumptions underpinning a program theory, developmental evaluation brings these issues much more to the fore.

According to Patton (2011) developmental evaluation is an approach that can deal with dynamic and volatile interactions, turbulent environments and unpredictable outcomes. Outcome patterns emerge from self-organisation among the interacting agents. This high degree of program and evaluation uncertainty implies the need for funders, program implementers and evaluation contractors to accept a similar high degree of flexibility in terms of contract agreements – something that my evaluation

work suggests is not generally forthcoming.

Westley and colleagues (2006) stress that developmental evaluation requires a long-term partnership between the evaluator and those engaged in the program. This, of course, may present a dilemma in terms of evaluation contracting and budget as highlighted in Chapter Three and in an example of evaluation using complexity ideas by Kremser (2011) (see below, section 4.7.1).

Developmental evaluation aims to provide rapid feedback as the program evolves, and opportunities to nurture exploration of ideas and for reflection (Westley, et al., 2006). It therefore appears similar to action research in that the process uses a cycle or spiral to pursue action (or change) and research at the same time:

> Developmental evaluation helps identify the dynamics and contextual factors that make the situation complex, then captures decisions made in the face of complexity, tracks their implications, feeds back data about what's emerging, and pushes for analysis and reflection to inform next steps, and then the cycle repeats (Patton, 2011 p30).

Both developmental evaluation and action research are therefore iterative and emergent processes. According to Patton (2011), while action research can be used as methodology for developmental evaluation, the distinguishing feature is that developmental evaluation focuses on program development, whereas action research focuses on the problem (Patton, 2011). However, the action research spiral as described by Kemmis and McTaggart (1988) shows how reflection from action research contributes to program development in the next iteration. Further, Patton suggests that theory-driven evaluation is 'top down' and participatory action research is 'bottom up', whereas developmental evaluation takes the middle ground, working in the space where evidence and local knowledge meet (Figure 4.5). However, it could be argued that theory-driven evaluation is only top down if the evaluator

comes in with a predetermined theory and, in fact, most theory-driven evaluators recommend working with stakeholders' beliefs and values to derive the program theory.

In Patton's model, local knowledge from innovation, adaptation and emergence arises from the microsystem of the local context. These ideas and practices may be expanded and taken up elsewhere. Best practice and principles of effectiveness arise from dissemination of models, evaluation and evidence and intersect with global or national macro systems. The middle ground is the space where top down and bottom up mix through developmental evaluation that generates large scale principles of effectiveness and, at the same time, nurtures local adaptation.

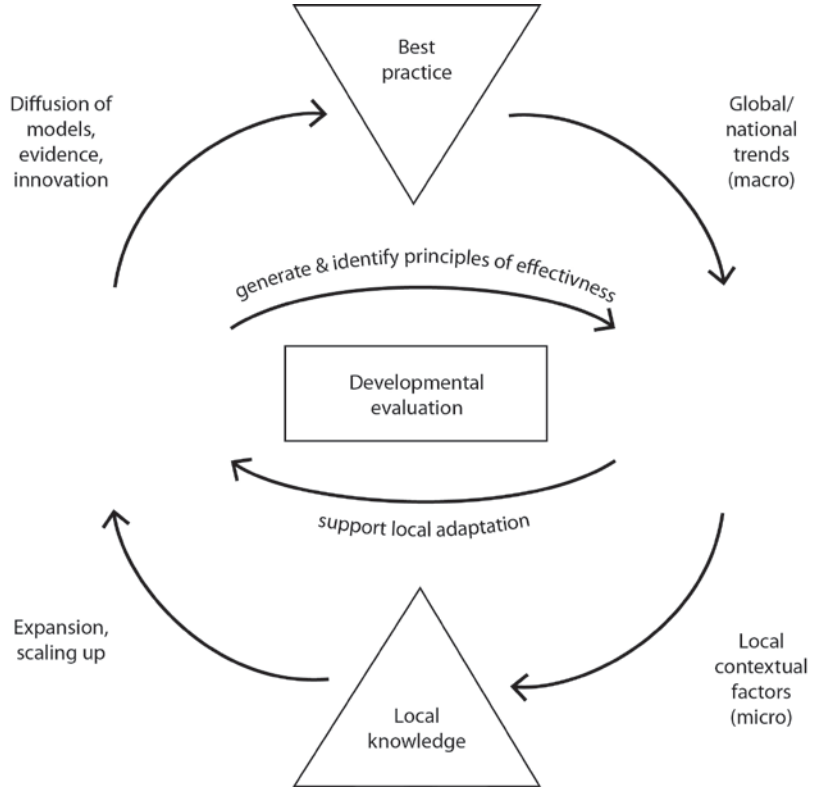

Figure 4.5 Developmental evaluation and the middle ground (based on Patton 2011 p180).

In summary, developmental evaluation aims to build on theory-driven and realist evaluation, with their focus on context, but adds a dynamic dimension in recognition of the notion that 'what works' is subject to change under complex conditions. This suggests that developmental evaluation can address a critique of program theory and program logic models that they tend to be linear and static and therefore not able to cope effectively with evaluation of complex community-based health promotion. The next section discusses a practical example of complexity thinking in evaluation of a health promotion program and also examines how complexity offers insights to the challenges presented in my published studies.

4.5 Use of complexity theory in health promotion evaluation

Having explored the synergies between complexity theory and health promotion evaluation this section considers application of this concept in practice. There are few practical example of complexity theory in use in the evaluation of health promotion programs. This section describes one example and then considers how complexity theory might address the evaluation challenges illustrated by my published studies and the possible insights from complexity theory for these evaluations.

4.6.1 Health promoting schools example

While it is too early for a body of knowledge on applying complexity theory to health promotion evaluation, Kremser (2011) provides a practical example of in the evaluation of a Health Promoting Schools initiative. Schools have been defined as 'social complex adaptive systems' (Keshavarz, Nutbeam, Rowling, & Khavarpour, 2010), and Kremser (2011) draws on complexity theory in evaluating how the school as an organisation influences the process of implementing a health promoting schools

approach. Kremser (2011) identifies structural determination (or self-organisation) and structural coupling (the engagement between systems that influences change in each) as key concepts in complexity theory and explores how these concepts influenced the implementation of the Health Promoting Schools initiative.

The evaluation used interviews and observations to build up an understanding of the system structures, with each classroom-based project seen as nesting into the school level intervention. One lesson from this example is that the approach requires frequent field visits and prolonged engagement due to the need to understand inter-relations between parts of the school and how these contribute to the whole school as a system. This investment of time also helped to build trust with the teachers and other stakeholders which led to higher quality data. Kremser (2011) found that the concept of structural coupling was relevant in analysing the processes and history of decision making that led to the choice to become a health promoting school. Structural coupling holds that how a system reacts to external influences is more dependent on the structure of the system than on the environment. This helps to explain why the school adopted a particular focus on limited health promotion topics and did not change its overall approach to health promotion despite joining the Health Promoting Schools network.

Structural determination suggests that what changes will occur in a system, and the change process, is determined by the existing structures. In this case, the tendency for primary school teachers to act autonomously in their classrooms limited the capacity of the school to act as a system to adopt the Health Promoting Schools approach. The implications drawn by Kremser (2011) are that health promotion interventions in social systems need to be designed with existing structures in mind

and that complex systems are unpredictable. Evaluations need to identify existing patterns of relationships, interactions and structures as these factors all have an impact on the implementation and achievement of the health promotion initiative.

4.6.2 How complexity theory could have contributed to my published studies

This section draws on complexity theory as outlined above to assess what contribution this could have made to my published evaluations. First, the studies are classified according to the Cynefin framework (Table 4.3). An overall placement is made along with classification of major components of the project.

Table 4.3 illustrates my overall placement of the studies into complicated or complex domains by considering the extent and diversity of components, stakeholders and interactions. Further analysis distinguishes between complicated and complex components of each study. As described earlier, complicated interventions or components are characterised by multiple objectives and stakeholders and unclear cause-effect pathways. Complex interventions have feedback loops, adaptations and multiple components that may act independently or interdependently, resulting in unpredictable cause-effect relationships. Thus, multiple stakeholders or components make a program complicated; complexity arises when the program implementation and outcomes are dependent on the interactions between these factors.

Next, Table 4.4 presents the challenges and potential benefits for my published studies from drawing on complexity theory and developmental evaluation approaches.

Table 4.3 Classification of studies into the Cynefin framework

Evaluation study	Main classification	Complicated components	Complex components
Eat Well SA	Complicated	Six components/sub-objectives Six working groups (including Evaluation Working Group) Multiple stakeholders with different perspectives Multiple activities Multiple sites for activities	Generative change Community setting Interactions and partnerships between different sectors
What makes for sustainable Healthy Cities initiatives?	Complex	Multiple stakeholders with different perspectives	Interactions and partnerships between different sectors High level of community control Community setting Political changes over timeframe of project High levels of emergence
Building an evidence base for community health	Complicated	Multiple, diverse programs and evaluation reports Practitioner skills and resources	Organisational cultures
Framework and tools	Complicated	Engaging with multiple stakeholders with different perspectives	
Healthy Ageing – Nutrition	Complex	10 diverse organisations with individual goals and strategies, different starting points and varying resource constraints	Highly dependent on project manager's relationships with participating organisations Organisational cultures Changing organisational operations during lifetime of project Political and funding changes Action research and emergence of goals and strategies

Table 4.4 How complexity theory and developmental evaluation could have benefitted my evaluation research

Evaluation/ research project	**Challenges**	**Potential benefit of complexity/ developmental theory**	**Different approach**
Eat Well SA	Coming in as evaluator to an established program is not ideal	More flexibility in program theory	More time developing (and getting stakeholder agreement on) the program theory and logic model. Recognising the complicated and complex components of the project and adjusting the evaluation to suit
What makes for sustainable Healthy Cities initiatives?	Lack of capacity to undertake systematic monitoring and evaluation of project from start	Increased monitoring of emergence and context	From the start, look for and document emergent ideas, interactions etc. Look for tension between these forces and the concept of sustainability
Building an evidence base for community health	Unable to draw generalised lessons from diverse evaluation reports	Generalised theory rather than standard program	Extract and articulate the mid-range program theories and test in a new context
Framework and tools	Tension between evaluation expertise and reflexive practice	Increased empowerment of practitioners	Engage with 'evidence' and 'knowledge', act more in the 'middle ground'
Healthy Ageing – Nutrition	Limited participation by stakeholder organisations	Increased stakeholder participation	Actively seek more feedback on findings and generalisable theories

In summary, the evaluations would have benefited from program theory that was more flexible and responsive to developments over time, and from looking for opportunities to articulate more mid-range theories in order to increase transferability of findings. Increased stakeholder engagement would have strengthened the evaluation designs and credibility of the findings.

4.6.3 Lessons from complexity theory for Healthy Ageing – Nutrition evaluation

Finally, this section considers how complexity theory could have been applied to the *Healthy Ageing – Nutrition* evaluation. The *Healthy Ageing – Nutrition* project contains aspects that could be described as simple, complicated or complex, while overall it represents a complex intervention. This overall complexity is demonstrated by the adaptive approaches undertaken to address the needs of ten very different participating organisations. These organisations all developed their own emerging objectives under the umbrella of the project aims. Partnerships and relationships were crucial to the success of the project and these were also emergent over time.

An ideal evaluation would consider all aspects of the program and choose an appropriate evaluation approach and method for each. Simple or ordered aspects of the project included the production of resources and the web-site. Evaluation questions could ask whether the resources and website were designed following identified good practice and about the number of hits on the website and the number of resources requested. Whether the resources prove useful and effective for the participating organisations is a more complicated evaluation question since the different organisations have different levels of need for information and opportunity to make use of the resources. The evaluation would need to assess if the resources were accessible and relevant in each context of use and whether they contributed to

any change in skills, behaviour or policy.

A complex aspect is that each organisation worked to its own action plan and implemented this plan in the context of organisational-specific enablers or constraints. So, some organisations relied heavily on volunteers while in others the workforce was predominantly health professionals. This is an example of the structural determinism described by Kremser (2011) above and illustrates the notion of complex systems as self-organising and dependent on history. The evaluation needs to be aware of and understand these contextual differences in assessing the extent to which organisations achieved their own objectives and the project aims, thus drawing on ideas of context evaluation rather than solely program evaluation.

The differences in the participating organisations also lead to complex aspects of the project. For the project manager, how organisations responded to the initial invitation to join the project was unpredictable, as was how they went on to be involved. Relationships played a key role, not only the relationship between the project manager and the organisational stakeholders but also relationship within each organisation. Thus the program illustrated uncertainty in how it would be implemented and was dynamic in terms of developing relationships and interactions. For this aspect the evaluation needs to document the initial conditions and monitor changes as they emerge. These changes may be directly related to the project, for example, introduction of a new policy, or indirectly related, such as a change in key personnel. My evaluation attempted to build rapport with the key stakeholders and gain an understanding of the context for implementation but this was limited by the evaluation resources and the capacity of organisations to commit time to this. Complexity is also apparent when considering the project overall, with its aim to

develop workforce capacity to improve the nutritional health of older people in South Australia. The project found little agreement or certainly about who makes up the relevant workforce, what capacity-building was needed and the best strategies to use. Following Hawe, Shiell and Riley's (2009) advice, overall systems level change is best captured by assessing to what extent change is embedded in each organisation, tracking changes in relationships and networks across the system, documenting changes to how resources are distributed and looking for activities that have been displaced in order to take on the new activities. For *Healthy Ageing – Nutrition*, the timeframe and resources were barriers to evaluating beyond the first of these.

4.7 A conceptual model of community-based health promotion evaluation

This thesis has considered the context for my published evaluation work and the challenges in evaluating community-based health promotion. I have argued that mainstream evaluation practice has to contend with demands for a linear, objective scientific approach that does not sit well with community-based approaches and health promotion principles and values. My evaluation work and the arising publications illustrate the negotiation and compromise needed in taking a more interpretive approach that can overcome some of the challenges inherent in evaluating community-based health promotion. Drawing on complexity and developmental evaluation concepts, I combine these with lessons from my work to construct a conceptual model that illustrates three phases of planning, implementation and evaluation of complex community-based health promotion initiatives (Figure 4.6.). The model is designed to address some of the issues for evaluation identified earlier. These issues often include a lack of well-developed program theory, the linear and static nature of logic models, a somewhat simplistic

and reductionist description of the intervention, problems with participation, and the context-contingent nature of implementation. Unintended and unanticipated outcomes can also be missed. The model therefore builds on theory-based evaluation by including program theory and program logic models but recognising that these should be subject to revision in the face of findings from evaluation of implementation process and outcomes. Participatory approaches and frequent feedback to support program development form part of the model. Developmental evaluation is drawn upon to help deal with complexity characteristics by acknowledging expert and lay input to the program theory and logic models, and by recognising the adaptive, dynamic and emergent nature of the initiative. Thus, the model more closely integrates program theory and complexity theory and illustrates a way for the theory and logic to be revised in response to evaluation findings. It also applies a specific health promotion lens.

The model illustrates the three phases of planning, implementation and evaluation. Health promotion principles and values feed into all these phases. The layout of the model is not intended to suggest that planning, implementation and evaluation are linear. In a complex system these interact and the arrows on the left hand side illustrate the two-way flow of information between each phase. Reflecting one of the lessons from this thesis, the model proposes that the evaluator, or evaluation team, should be engaged at all phases rather than being brought in when the initiative is well-established or towards the end of its life.

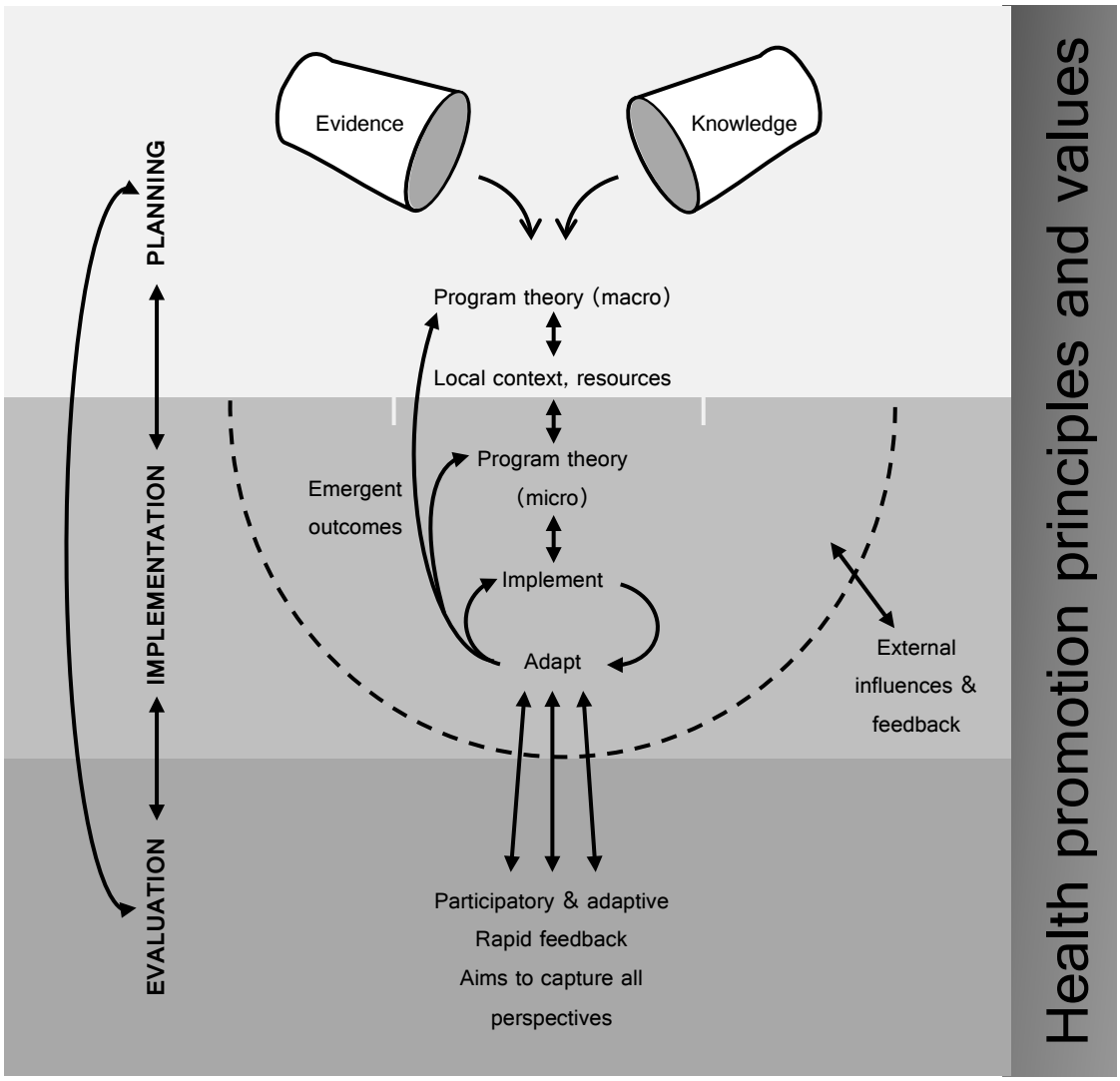

Figure 4.6 Planning, Implementation and Evaluation conceptual model

The planning phase starts with input from two 'buckets' of information. The evidence bucket contains evidence from academic literature, accepted best practice and findings from other research and evaluation. The knowledge bucket content is mainly practitioner and lay wisdom about what has worked before or might be expected to work in this context. These two sources of information are then used to produce a macro program theory and logic model, including predicted outcomes, to provide overall guidance to program development. Program planners, funders, managers and evaluators should, ideally, work collaboratively to produce the

program theory. The program theory at this stage is an overarching model of how the program is expected to work, and what outcomes are likely to be achieved, based on previous evidence and knowledge. The program theory then needs to be adapted to the local situation and this requires two filters: context and resources. Consideration of the local context and community involves, for example, identification of the population group for the program, the geographical area, the needs and strengths of the community, experience with previous initiatives, local stakeholders and decision makers, and social, economic and political factors. The resource filter provides a 'reality check' in terms of financial and human resources for the program, but also identification of the priority activities and evaluation questions. The capacity, skills and resources of practitioners and community members are also important factors. Applying these filters leads to the development of a more locally specific, micro, program theory or possibly a series of theories. A localised program logic model(s) would identify the links between program components and expected outcomes taking into account the available resources and the enablers and constraints in the local environment or setting. Ideally, in addition to the group developing the macro program theory, a broader range of stakeholders would be involved in developing the specific theories, including the practitioners who will be delivering the program, community representatives or leaders, and people from the agencies and groups who will be partners to the initiative. Thus, in the planning phase, formal evidence and the lay wisdom of stakeholders are brought together, as described by Patton (2011) in developmental evaluation. A strength of the model is that it then combines these to make the program theory explicit and to draw out the program logic model in a collaborative exercise. Further, the model allows for recognition of the impact on the program theory of local contextual factors (the micro) and likely resource constraints

as identified in Chapter Three. The result of this filtering process is that a context-specific and realistic program logic model can feed in to the implementation stage.

Program activity takes place in a permeable 'implementation space'. The model illustrates that the implementation of complex interventions does not normally follow a linear path. Implementation and adaptation flow in a cyclical manner and this is influenced by internal and external factors, history, networks, relationships and feedback loops as suggested by complexity theory. The program theory and components of the logic model are continuously adapted to take these emergent issues into account and, in this way, the critique of program logic models as too static and linear is addressed. The evaluator's role during this stage is to monitor the internal and external influences and provide rapid feedback to stakeholders so that the initiative can be adapted in response to emergent issues. This process draws on developmental evaluation in that it encourages the evaluator to engage with the implementation in order to provide timely feedback and support the program as it develops.

Evaluation flows from, and interacts with, the implementation space. Evaluation is underpinned by health promotion principles of participation and empowerment so that as far as possible, it is participatory and adaptive, provides rapid feedback to stakeholders and aims to capture all perspectives. Reflexive evaluators therefore need to ask themselves how, and to what extent, the evaluation process: encourages and supports participation by stakeholders; captures all perspectives; adapts to changing system dynamics; monitors interactions; and provides rapid feedback. These reflexive questions are answered, or least uncertainty about them is reduced, by documenting, analysing and interpreting: actions, decisions, adaptations; initial

conditions and changes in internal and external environments; changes in networks, relationships, interactions and why these occur; unanticipated events; conflict, disagreement and uncertainties. Findings from the evaluation are context-specific in that they relate to the micro program theory as it is adapted during implementation. However, by examining findings against both the macro and micro program theories, overarching 'principles of effectiveness' (Patton, 2011) may be identified. That is, established or emerging evidence of good practice may be supported or countered. The ability to identify principles of effectiveness may increase as findings from similar programs are built up through repeated application of the model.

Applying the Planning, Implementation and Evaluation Model to *Healthy Ageing – Nutrition*

Figure 4.7 provides a practical example of how the model might be used, drawing on the evaluation of *Healthy Ageing – Nutrition.* The *Healthy Ageing – Nutrition* project had a number of complex components including: interactive relationships, uncertainty about defining the workforce, organisational cultures, diversity and ongoing changes in organisations, power differences within and between organisations, changing political and economic contexts. These led to unpredictability about processes and outcomes and, in retrospect, the evaluation could have provided more detailed and rapid feedback to the stakeholders.

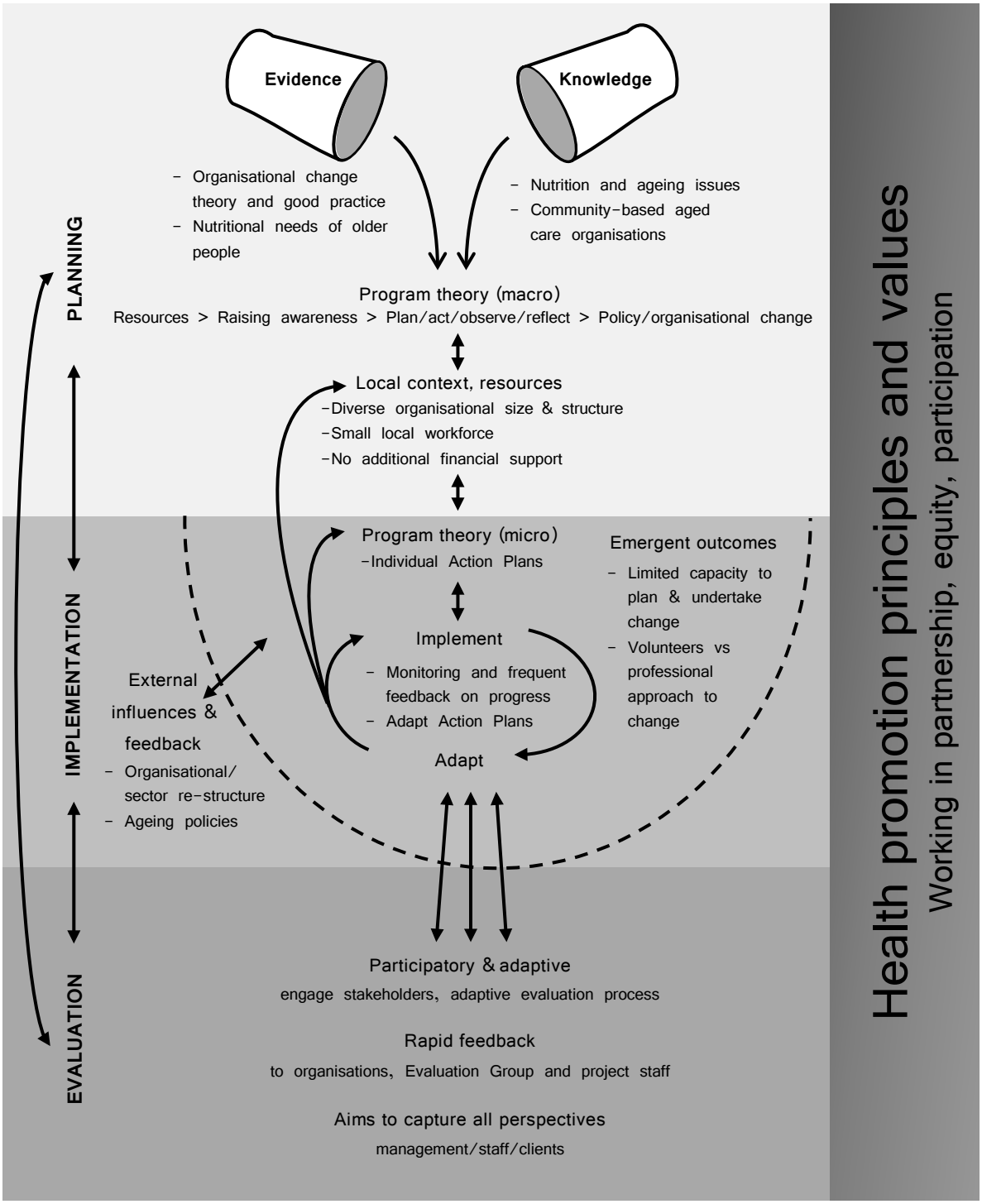

Figure 4.7 Application of Planning, Implementation and Evaluation model to Healthy Ageing – Nutrition

While some of the activities above were undertaken, applying the *Planning, Implementation and Evaluation* model suggests that more focus on the following

changes to practice would have been beneficial.

In the planning phase:

Identifying the complex and complicated components of the project would have enabled planning for the evaluation of these different components to be more effective. It would have alerted the Evaluation Group to the uncertainties in how the program might unfold and acted as a prompt to look for where complexity issues might have an impact on implementation.

While there was some documentation of existing evidence and good practice in enabling organisational change in the healthy ageing community-based services context, this could have been strengthened. This would have allowed existing evidence to play a larger role in the planning and implementation of the program and provided a framework for development of a program theory and program logic model to guide the evaluation. Documenting the practice knowledge of the project manager and key stakeholders about stimulating and supporting organisational change in the healthy ageing community-based services context would also have added to the understanding of the program's underlying theory.

Applying the *Planning, Implementation and Evaluation* model would have led to combining evidence and practice knowledge to develop a program theory and program logic model for the *Healthy Ageing – Nutrition* intervention. The implication from complexity theory is that the macro program logic model needs to be adapted for the diverse organisations engaged in the program. Thus, the macro program logic model would be filtered through the local context for each participating organisation and the resources available to it. For example, organisations came to the program with varying levels of financial and human

resources and capacity for change. The willingness and capacity of stakeholders to draw on evidence and practice to develop program logic models also varied.

These changes in the planning phase all require the evaluator to have input at the early planning stages of the program and this is likely to increase the level of evaluation resources required. It also needs the funder or program auspice to be supportive and cognisant of the need for evaluation input at this early stage of planning.

In the implementation phase:

Changes in the implementation phase would include more systematic monitoring and assessment of factors influencing implementation in order to gain an understanding of how these factors shape the context, and how this changing context reacts with the program and its theory. Complexity theory highlights the need to document emergent interactions, feedback loops, adaptations of the planned program and activities rather than focussing solely on processes and impacts expected from the original program logic model. In practice, resource constraints on the *Healthy Ageing – Nutrition* evaluation limited capacity to undertake a more comprehensive analysis of the adaptations and emerging issues during the implementation phase. For example, it became apparent that different approaches to stimulating change were required when dealing with a volunteer workforce compared to a professional grouping. One participating organisation was undergoing large structural change and this had implications for the way the program was implemented in this case.

To enhance participatory, developmental evaluation, the *Planning, Implementation and Evaluation* model suggests the need to provide frequent feedback to stakeholders, particularly participating organisations. This engagement with

participating organisations requires a high degree of program ownership and investment by participants and capacity for this in the participating organisations varied.

The changes suggested during the implementation phase require evaluation resources to be used to monitor and assess systematically the environmental changes that potentially have an impact on implementation. Participants and stakeholders would also need an interest and understanding of evaluation and the capacity to engage in this aspect of the project.

In the evaluation phase:

Changes in this phase would mean that the evaluator analyses the influence of the internal and external environment on implementation and subsequent impacts from the project. Emerging issues and early findings would be continuously negotiated with stakeholders in order to provide proactive and timely feedback. In the *Planning, Implementation and Evaluation* model, environmental analysis is a critical component of the evaluation and should assist the program to remain relevant in a changing context. This analysis would then be drawn upon to adapt the evaluation process to the unfolding and emergent program. This should enhance the usefulness of the evaluation as it reflects the actual events rather than only what was predicted. For example, in the *Healthy Ageing – Nutrition* project of collaborative links between participating organisations were made that had not been anticipated but were likely to continue after the life of the project.

The *Planning, Implementation and Evaluation* model suggests that, as far as possible, the evaluation should aim to engage all stakeholders so as to gain a broad perspective and give voice to those who may not otherwise be heard. The *Healthy*

Ageing – Nutrition evaluation involved leaders of the participating organisations and some other stakeholders but no input was sought directly from the workforce or clients. This was outside the scope of the evaluation but would have added to the comprehensiveness of the evaluation findings.

Developmental evaluation and complexity theory suggest that it is important to ask reflexive questions as a way to assess the quality of the evaluation and the evaluator's own professional development. Considering the reflexive questions (outlined on page 163) would have served as a prompt to examine the overall quality and effectiveness of the *Healthy Ageing – Nutrition* evaluation process. For example, was participation by all stakeholders (including participating organisations) supported? Were all perspectives captured? (only organisation leaders were interviewed, not the wider workforce or clients). Was evaluation adapted to changing system dynamics? Were interactions and power differences monitored? Was rapid feedback provided?

Finally, the *Planning, Implementation and Evaluation* model suggests identifying principles of effectiveness that may be transferable to other contexts. These principles can inform future evaluations of similar programs by feeding into their program theories. While analysis of the evaluation data for *Healthy Ageing – Nutrition* did lead to a model of enhancers and barriers to making changes in the participating organisations, this would have been strengthened, and transferability increased, by going back to refine the original program logic model in light of the evaluation findings. However, opportunity for this is constrained in time-limited projects such as *Healthy Ageing – Nutrition.*

4.8 Chapter summary and conclusions

Community-based health promotion initiatives present many challenges for evaluators. Program theory and logic models can help to articulate the theoretical underpinnings of the program and how it is expected to work. However, these approaches risk imposing linearity and certainty on programs that are occurring adaptively in a dynamic setting. This chapter has presented complexity theory as a potential new approach for the evaluation of community-based health promotion initiatives. Complexity theory resonates with many of the issues raised in evaluation of health promotion: dynamic interactions and changing contexts driving uncertainty about processes and outcomes. Recognition of the complexity of an intervention and its setting leads to approaches such as developmental evaluation which emphasise the need for evaluation to be responsive to the program and its changing context and to work in partnership with program stakeholders.

An example of a health promotion intervention evaluation drawing on complexity theory suggests this is a useful framework for designing an evaluation. It does, however, require a high investment of time and effort from the evaluator and the stakeholders to build trust and understanding. Similarly, the example from *Healthy Ageing – Nutrition* highlights the need for a high level of engagement from participant organisations, and levels of resourcing that enable the evaluation to extend beyond assessment of short-term, localised change to higher order system change over time.

This chapter has drawn on the lessons from my practice and more recent approaches to evaluation to construct the Planning, Implementation and Evaluation model as a way to conceptualise reflexive evaluation of community-based health promotion

initiatives. This contributes to Research Question Four:

4. What are the overall lessons from the evaluation practice presented in my publications and how do they inform new approaches to evaluation of community-based health promotion initiatives?

The model builds on developmental evaluation by more closely integrating program theory and complexity theory and by applying a specific health promotion lens. In this way the evaluation may increase understanding of the unique actors, power differences, conflict and environmental context of the health promotion initiative and so unravel some of the complexities towards a situation of more knowable complication. This is important because evaluations of interventions that are complicated, rather than complex, are more able to test program theory and predict cause-effect linkages.

In terms of the issue of transferability of findings, a further research question might be asked: to what extent can the learning from evaluation of this initiative be applied in other situations? This means trying to identify the principles and assumptions underlying the program theory that contribute to effectiveness and that might be applicable in a different context. These principles for effectiveness could then be added to the evidence bucket if sufficiently robust, or to the knowledge bucket if not yet formally established, when planning for another initiative. In this way, cumulative learning could be used to build up knowledge about what works across a range of contexts. The model, then, aims to contribute to reducing complexity and to the prediction of effective practice principles that might be applicable in a different setting or context.

CHAPTER FIVE: DISCUSSION AND CONCLUSIONS

5.1 Introduction

This thesis has drawn on health promotion and evaluation literature and my published evaluation work to argue that the nature of community-based health promotion makes evaluation difficult. Partly this is due to the contested understanding of health promotion itself and partly due to the dominance of a positivist paradigm in research and evaluation. Although evaluation theory from the late 1970s has evolved to include more interpretive approaches, mainstream evaluation practice still has to contend with demands for a linear, objective scientific approach that does not sit well with community-based health promotion. My evaluation work and the arising publications illustrate the struggles and compromises in taking a more interpretive approach. I argue that using program theory and complexity theory together can help evaluators to overcome some of the challenges in undertaking evaluation in complex health promotion settings.

This chapter briefly reviews the health promotion and evaluation developments identified earlier in the thesis and how these relate to my evaluation work. Strengths and limitations of the thesis are described. I then discuss responses to my four research questions by drawing on the evaluation and health promotion literature and on my published papers as presented in preceding chapters of this thesis. Finally, I revisit the Planning, Implementation and Evaluation model presented in Chapter Four and draw some conclusions about future directions for evaluation of community-based health promotion initiatives.

5.2 Health promotion and evaluation developments

Over the last four decades debates have continued about the effectiveness of the two broad approaches to health promotion practice. The first, based on social marketing and education theory, leads to health promotion initiatives such as mass media campaigns to raise awareness of a health issue and educational materials aimed at increasing knowledge and skills in this area. These activities are expected to produce behaviour change to increase healthy lifestyles and decrease individual risk factors and hence improved health. The second approach uses socio-environmental theories based on social action and intersectoral collaboration, in order to facilitate behaviour change through changes in health determinants and in increasing supportive environments for health. These different approaches and theories are not mutually exclusive but are often used together to support health promotion practice.

Evaluation theory and practice have also evolved over time. A range of evaluation models has developed with more recent approaches focussing on uncovering the theory or program logic that underpins the steps within a program and then finding appropriate ways to test this theory. The realist approach as promulgated by Pawson and Tilley (1997) emphasises the importance of including the context and mechanisms influencing behaviour change. This progression suggests that evaluators are recognising the dynamic interactions and networks at work in a community-based intervention and are searching for evaluation approaches that can understand how these contribute to the achievement of outcomes. However, as my publications show, many challenges remain in finding evaluation approaches that can deal with these complex characteristics. Complexity theory and developmental evaluation offer new approaches and show promise in progressing evaluation of community-based complex interventions.

My standpoint is that health promotion is underpinned by a set of principles and values. Much health promotion literature suggests that evaluation approaches should be synergistic with these principles and values (Judd, Frankish, & Moulton, 2001; Kelly, et al., 2007; McQueen, 2007; Poland, 1996a; Rootman, et al., 2001; Tones & Green, 2004; Wass, 2000). In this view evaluation practice should, ideally, aim to be participatory and empowering for all stakeholders, encompass assessment of partnerships and multi-sectoral practice, and the extent of community/individual participation. Health promotion evaluation should monitor and document changes in health equity, and individual and organisational capacity for health promotion, in addition to health outcomes for individuals and populations.

5.3 Strengths and limitations of the thesis

The review of evaluation and health promotion literature has helped in the identification of underlying theory in my papers and to link this to my developing evaluation practice. My journey as an evaluator reflects, to some extent, the development of evaluation approaches over time. The thesis contributes to future evaluation development by analysing the potential role of developmental evaluation and complexity theory in aiding the evaluation of complex community-based health promotion initiatives.

My publications were not written with their use in this thesis in mind. They represent dissemination of evaluation methods and findings for a mixed audience of researchers, policy-makers and health promotion practitioners. The papers describe and discuss relatively small scale evaluations of relatively small scale community-based health promotion initiatives. As such, there was limited scope for exploration of the theoretical underpinnings of the programs and the papers focus on practice.

Nonetheless, theory is implicit in the works and has been extrapolated in Chapter Three. It should be noted that one paper *Building an evidence base for community health: a review of the quality of program evaluations* reports research investigating programs and activities conducted in metropolitan South Australian community health services. Not all these programs would fit my criteria for community-based health promotion as used in this thesis but they represent the reported range of evaluation activity carried out by community health services at that time.

A further limitation concerns the proposed conceptual model. This remains a theoretical construct and has not been tested in practice. I have applied the model retrospectively to the *Healthy – Ageing Nutrition* evaluation as a first test of its utility. An area for further research would be to test more rigorously the usefulness of the model in a variety of real-time health promotion evaluations.

5.4 Addressing the research questions

The aim of the thesis is to describe, assess and contribute to addressing the theoretical and practical dilemmas arising from the evaluation research presented in my publications. This section discusses the findings from the study in addressing my research questions.

5.4.1 Health promotion and evaluation context and influence my evaluation work

Chapter Two presented a review of health promotion and evaluation literature in order to set the context for this thesis. Health promotion is contested as a discipline (Tremblay & Richard, 2011) and is underpinned by multiple theories (Nutbeam, et al., 2010). A continuum of activities is apparent (Labonte, 1992), however, the tension between individual lifestyle education and the structural/environmental approach has continued since selective primary health care emerged from differing

interpretation of the Alma Ata Declaration and the Ottawa Charter (World Health Organization, 1986a). Health promotion principles drawn from the Ottawa Charter stress the need for multiple strategies, community empowerment and participation. In practice, however, funders and policy makers tend to focus on social marketing and individual behaviour change interventions in preference to more population-based socio-environmental actions (Baum 2011). In my evaluation work this is illustrated by the high level of interest shown by those in power, such as program designers and funders, in the reach and scope of educational resources and an expectation that this will lead to behaviour change. At least in part this may be due to the apparent ease of implementing and assessing these former approaches compared to the political challenges of targeting structural changes (Peersman, 2001). It is also likely to reflect the dominance of the objective, scientific paradigm that gives power to clinical and medical understandings of health and illness. The Healthy Cities, and to some extent, *Healthy Ageing – Nutrition* evaluations best reflect the effort to tackle structural change.

Health promotion arose from a medical paradigm and remains linked with health and medical services (Green and Tones 2010). It is no surprise then that the introduction of evidence-based medicine soon led to calls for evidence-based health promotion and a surge of interest in measuring the health outcomes from health promotion activities. This has led to something of a paradox in that while many health promotion interventions claim to be founded on a broader social model of health, research and evaluation is still largely expected to follow objective, scientific methodologies. This meant, for example, that 'selling' the concept of generative evaluation in *Eat Well SA* required persuading the program overseers to relinquish control over aspects of the evaluation and accept some uncertainty. However,

evaluation of community-based health promotion does have to deal with much complexity, for example, permeable boundaries, diversity of practice and settings, multiple interactions and stakeholders, and the impact of the social and economic determinants of health (Dooris, 2005; Green, et al., 2000; Tones & Green, 2004).

In the decade preceding my publications, theory-based evaluation was becoming more dominant in social programs, with a change of focus from the method of measurement to identifying underlying program theory and articulating logic models (Birkmayer & Weiss, 2000; Rogers, et al., 2000; Rossi, et al., 2004). By 1997, realistic evaluation (Pawson & Tilley, 1997) was beginning to be established with early examples in criminology. Evaluation is now more likely to be driven by theory rather than the choice of a particular method of data collection (Chen, 1990). Context is recognised as critical as are the many different stakeholder perspectives and the dynamic interactions and networks at work in a community-based intervention. Reflecting on my evaluation work suggests that commissioners of health promotion evaluations are not yet fully convinced of the need to recognise complexity and for evaluation to probe the significance of context and interactions on how a program is implemented and contributes to change.

The contextual influences on my evaluation work illustrate that, at the time, a positivist paradigm was still driving expectations for evaluation, and the interpretive approach was viewed by evaluation commissioners as somewhat marginal. My evaluation work also had to contend with a mainstream view of health promotion as social marketing, predominantly concerned with changing individual behaviours. These factors meant that my evaluation work included advocating for new approaches and for recognition of a broader understanding of health promotion to

include efforts towards structural and social change. These new approaches appreciate the complexity of community-based health promotion and the need for more appropriate evaluations.

5.4.2 Evaluation developments including the changing role of the evaluator

The development of theory-based approaches to evaluation has provided a useful framework for community-based health promotion evaluations (see, for example evaluations of *Have a Heart, Paisley* and *Starting Well* (Mackenzie & Blamey, 2005) and *Health Action Zones* evaluation (Judge & Bauld, 2001) although both sets of authors also identify challenges in practice. Nonetheless, these approaches offer a way to consider issues of causality in a complex environment where controlled experiments are neither feasible nor ethical. Indeed, understanding the context, rather than controlling it, is considered crucial to conducting a theory-based evaluation.

My evaluation practice, as reflected in the five articles forming the basis of this thesis, has moved from the 'black box' approach with a focus on process and impact measurement to using stakeholder perspectives to articulate program theory and increase understanding of the importance of context and how this influences what works for whom. The *Eat Well SA* evaluation demonstrates my early thinking about this, with traditional process and impact evaluation being supplemented by the generative component that was used to develop a program logic model for the capacity building aspects of the project.

The role of the evaluator has expanded from value-free technical measurement to theory development, negotiation and partnership, and my role has similarly moved along this continuum, albeit constrained by the evaluation context and budget. Thus, I have developed a greater understanding of the importance of relationship building

and effective communication with program managers and other stakeholders. This shift is best illustrated by the *Healthy Ageing – Nutrition* evaluation where the action research approach of the project meant that it was beneficial to act as a partner to the project and to work closely with the program manager. This was aided by the strong element of reflexive practice brought by the program manager and the committed interest of the participant organisations.

As evaluation has come to be accepted as a more values-based enterprise, the need for health promotion evaluation to reflect the principles of health promotion has also become more apparent (Green & Tones, 2010; Potvin & McQueen, 2009; Rootman, et al., 2001; Wass, 2000). In my evaluation work, I have strengthened my focus on health promotion principles as aspects of evaluation practice, to incorporate assessment of collaboration, partnerships and participation, and equity impacts wherever possible. While these aspects may not always be explicit in the program design, a developmental evaluation approach can be a way to develop a focus on these as the program develops. Program managers and practitioners can be alerted to the importance of the inclusion of health promotion principles through the identification of their place in the program logic model and by the evaluation including an assessment of their application. The *Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion* paper, for example, provides a rationale and tools for evaluation of some health promotion principles. Reflecting on my experience, funders and decision makers are also increasingly interested in evaluation of these program components.

These developments in evaluation theory and practice do not mean that all evaluation has followed this path. Rather than replacing earlier approaches, evaluation is a

mixed field of endeavour with positivist and interpretive paradigms co-existing. Thus there are multiple evaluation approaches, giving the opportunity to choose the most effective and appropriate. This means that evaluators should be prepared for multiple roles, including understanding the changing context and dynamic interactions at play in community-based interventions.

5.4.3 Theoretical, methodological and practical challenges in conducting community-based health promotion evaluations

The evaluations reported on in my publications illustrate some of the theoretical, methodological and practical dilemmas that influence the extent to which an ideal evaluation can be conducted. These include evaluation design issues and the capacity and resources available for evaluation.

Community-based health promotion initiatives are often dynamic, complex systems. In terms of evaluation design, experimental methods are rarely appropriate or realistic in this context. It is difficult to measure change in health status in the long timeframe often needed for health promotion outcomes, and, even if change is seen, causal links are problematic. A more realistic approach is to focus on shorter term outcomes such as knowledge, skills and behaviour change and draw on theory or evidence from other programs about the likelihood of these contributing to improved health. The links between these short and longer term outcomes can usefully be expressed in a program logic model. Particularly for more complex interventions, such as Healthy Cities, the program theory and logic model is likely to be emergent and subject to change as the context and players interact and change. This means the program logic cannot be static but should be flexible to enable adaptation to new situations. Since community-based health promotion programs operate in an open and dynamic environment, it is important to document change in the social,

economic and political context and assess the influence of this on the way an initiative unfolds and its outcomes. The notion of complexity suggests that it not possible to predict with any certainty the outcomes from complex interventions but that evaluation should identify initial conditions and monitor what emerges, facilitate reflexive practice and embed evaluation in the intervention. The evaluation should act to monitor, document and feedback changes in order to drive the development of the program's implementation (Patton, 2011). However, this approach is likely to be challenging for commissioners of evaluations to accept since it rejects program implementation fidelity and means that the evaluation framework and time-lines need to be flexible. It also requires new skills from evaluators such as recognising and incorporating multiple perspectives and dealing with uncertainties.

I have taken the standpoint that empowerment of individuals and communities is an underpinning principle of health promotion theory and practice. This means that evaluation of health promotion programs should also be empowering for the program recipients and other stakeholders. This presents a challenge since, as we have seen in Chapter Two, health promotion theory and practice are contested. Further, taking an empowerment approach requires those with power in the program (including evaluators) to be willing to share power and to be open to different perspectives. Implications from this are that the health promotion program and its evaluation design need to be flexible in order to allow for, and respond to, stakeholder input; the evaluator needs to take on a negotiator or mediator role in order to bring different interests to a workable consensus and stakeholders need the skills, resources and interest to be engaged meaningfully in the evaluation. *Building an evidence base for community health* and *Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion* both aimed

to assist in building evaluation capacity in health promotion practitioners.

We have seen that health promotion activities tend to be poorly funded compared to care and treatment and this difference extends to resources for research and evaluation (Minkler, Glover Blackwell, Thompson, & Tamir, 2003). A further resource issue is that health promotion programs and their evaluations suffer from short-term funding (as noted with *Eat Well SA* and *Healthy Ageing – Nutrition*). The capacity of health promotion practitioners and services to undertake or contribute to evaluation is limited and there is a need for evaluation support and resources. These could be provided from inside the organisation or through a consultancy service such as SACHRU. Since there is limited value in one-off evaluations, particularly of small scale programs, an institutional support system would enable consolidation of findings and a cumulative approach to knowledge about health promotion evaluation. This would assist with the ongoing challenge for health promotion evaluation design – the tension between the need for generalisable knowledge about good practice and the context specific setting for most programs.

These considerable challenges to evaluation mean that there is likely to be ongoing debate about how community-based initiatives should be evaluated and what mechanisms and outcomes should be the focus. Stakeholders to such programs need to engage in this debate, along with evaluators, so that health promotion can increase its standing as a well-researched strategy to improve health.

5.4.4 Overall lessons from my published evaluations and how they inform new approaches to evaluation of community-based health promotion initiatives

My studies have shown that community-based health promotion initiatives are often complex interventions set in a complex environment. The characteristics of these

interventions reflect complexity theory in a number of ways. The intervention acts through relationships, networks and interactions which are not predictable and there are multiple stakeholders with different perspectives, interests and power. There may be several, possibly competing, theories about how the intervention is expected to work and these theories are often implicit or poorly articulated. There are usually multiple components and strategies making up the program and these are likely to be interacting to be produce outcomes that are more than the sum of individual parts. In dealing with social issues, disagreement about the 'problem' and what to do about it, along with uncertainty about how to produce the desired outcome, is likely to be high. These complexity characteristics mean that patterns of behaviour cannot be predicted and observed patterns may not repeat. The adaptive and emergent nature of the initiative also means that evaluation may itself influence the processes and outcomes.

Understanding community-based health promotion as complex interventions in a complex environment leads to a number of implications for evaluation. In planning and designing the evaluation, the evaluator needs to negotiate and articulate the underpinning program theories and establish flexible program logic models that allow for responsiveness to feedback and adaptive development as the program unfolds. This process should draw on lay wisdom and local knowledge, as well as formal theory and evidence, in order to ensure that the local context is considered as a critical influence on how the program will unfold. All my published works illustrate a blending of academic and lay input. It is also important to document the starting context and how this changes over time. The Healthy Cities paper provides an example of conditions at the start of a program contributing to sustainability. In choosing data collection methods, the evaluator should be prepared to use multiple,

mixed methods that are most appropriate for the data needed to address the evaluation questions and to test the articulated theories. As far as possible, participatory approaches to evaluation can be facilitated by engaging with practitioners and program recipients in the planning, design and conduct of the evaluation.

During the evaluation, the evaluator will need to track actions and decision-making processes, in order to provide rapid feedback to inform development of the program and to enlighten disagreements and uncertainties. This feedback can also encourage reflective practice in practitioners and other stakeholders. While the program logic model will predict some outcomes, complexity theory suggests that, in reality, outcomes may not be predictable and the evaluation should also look for unanticipated events and outcomes as they emerge. This means the evaluation should be alert to what actually happens and what changes occur in individuals, communities and organisations. All my evaluation studies provide examples where this approach was taken. The evaluation can also take into account issues of participation and empowerment by documenting the perspectives, experiences and outcomes for different groups involved in the program and by tracking changes in interactions, distribution of resources, and power.

These lessons, taken together, suggest an expanded and multi-layered role for evaluation and for evaluators. Complexity thinking makes clear the interconnections and feedback loops between planning, implementation and evaluation that the evaluator needs to document. It also means that the evaluator should be prepared for uncertainty, emergent and unexpected outcomes, and dynamic interactions with the environment and between stakeholders. Taking a complexity approach requires

flexibility in planning and budgeting for evaluation. Thus, another role for evaluators is to educate evaluation commissioners about the need to accept more flexibility in contract agreements. However, the benefit should be increased understanding about effective community-based health promotion.

5.5 Conclusion

Health systems around the world appear to be moving to a transition phase between ordered/knowable and unordered/unknowable and even to the 'edge of chaos'. Reasons for this include the global financial crisis, the 'epidemic' of chronic disease, re-emergence of some infectious diseases, growing health inequalities and inequity, the threat of climate change and increasing natural disasters (World Health Organization, 2008). However, complexity theory suggests that transformational change is blocked by history and feedback loops (which lead to governments being locked into doing what they have always done), while critical theory suggests that change is impeded by the power of dominant ideologies (for example, positivism, biomedicine) and groups in society (for example, the medical profession). On a more positive note, the edge of chaos has been described as the space where innovation and creativity come to the fore. Both health promotion and evaluation face these opportunities and threats.

While Hawe and colleagues (2009) argue that it is the setting that is complex rather than the intervention, I argue that many community-based health promotion initiatives may be thought of as complex interventions set in complex settings. Communities and organisations can be considered as systems, with interactions and different roles for actors (Tones & Green, 2004). Community-based health promotion is made complex by the interplay of biological, behavioural, social and

environmental factors, the interacting components and self-organisation, social networks and relationships, multiple layers and diverse actors. If health promotion activities carried out in community settings are understood as complex, this influences the type of evaluation that can be conducted and the evaluation questions that can be asked.

This adaptive nature of many health promotion programs means that evaluations are context-contingent and dependent on an understanding of the internal and external environment. So, rather than external, objective evaluators, evaluators need to adopt a more reflexive practice. The conceptual model proposed in this thesis provides a framework for evaluation practice drawing on complexity theory and developmental evaluation. It suggests that formal and informal knowledge about the program and context should feed into development of a program theory and logic model. These are not definitive but are revisited as the program unfolds. Feedback loops linking community context and evaluation findings are used to adapt the program's implementation. Since the program is continually developed in response to contextual factors, evaluation needs to monitor this context and build evidence of effectiveness of the underlying theory or principles, rather than standardised implementation of program activities. Thus, the model is designed to assist evaluators in the process of evaluating complex interventions by supporting reflexive practice that can contribute to unpacking complexity elements and making these more 'knowable'.

Bringing together insights from theory-based evaluation and complexity theory can contribute to addressing some of the challenges for community-based health promotion evaluation. By recognising and responding to changing contexts and

emerging outcomes, providing rapid feedback and facilitating reflexive practice, evaluation is better able to cope with complexity. Drawing on complexity will enable the evaluation process to gain a better understanding of the influence of context and other implementation factors. This will contribute to building cumulative evidence and knowledge in order to identify the foundation principles of effectiveness that may be transferable to a new situation.

REFERENCES

Abma, T. A. (2005). Responsive Evaluation: Its meaning and special contribution to health promotion. *Evaluation and Program Planning, 28*, 279-289.

Alvaro, C., Jackson, L. A., Kirk, S., McHugh, T. L., Hughes, J., Chirop, A., & Lyons, R. F. (2010). Moving governmental policies beyond a focus on individual lifestyle: some insights from complexity and critical theories. *Health Promotion International, 26*(1), 91-99.

Antonovsky, A. (1996). The salutogenic model as a theory to guide health promotion. *Health Promotion International, 11*(1), 11-18.

Argyris, C. (1982). *Reasoning, Learning and Action*. San Francisco: Jossey-Bass.

Ashton, J., Grey, P., & K, B. (1986). Healthy Cities — WHO's New Public Health initiative. *Health Promotion International, 1*(3), 319-324.

Baum, F. (1995). Researching Public Health: Behind the qualitative-quantitative methodological debate. *Social Science & Medicine, 40*(4), 459-468.

Baum, F. (1998). Measuring effectiveness in community-based health promotion. In J. K. Davies & G. Macdonald (Eds.), *Quality, Evidence and Effectiveness in Health Promotion: Striving for certainties*. London: Routledge.

Baum, F. (2000). *The Effectiveness of Community-based Health Promotion in Healthy Cities Programmes:Issues of measurement and evaluation*. Paper presented at the International Conference on Healthy Cities and Urban Policy Research, Tokyo, Japan.

Baum, F. (2002). *The New Public Health* (2nd ed.). Melbourne: Oxford University Press.

Baum, F. (2003). The Effectiveness of Community-Based Health Promotion in Healthy Cities Programmes. In T. Takehito (Ed.), *Healthy Cities and Urban Policy Research* London: Spon Press.

Baum, F. (2011). From Norm to Eric: Avoiding lifestyle drift in Australian health policy. *Australian and New Zealand Journal of Public Health, 35*(5), 404-405.

Baum, F., & Cooke, R. (1992). Healthy Cities Australia: the evaluation of the pilot project in Noarlunga, South Australia. *Health Promotion International, 7*(3), 181-193.

Baum, F., Duffy, J., & Jolley, G. (Eds.). (2003). *Investing in Community Health - Finding the Evidence for Effectiveness*. Adelaide: South Australian Community Health Research Unit, Flinders University.

Baum, F., Jolley, G., & Bament, D. (2001). Healthy Cities Evaluation Framework Testing Project. Evaluation of a Healthy Cities Initiative: Noarlunga Community Action on Drugs. Adelaide: SA Community Health Research Unit, Flinders University.

Bickman, L. (1987). The function of program theory. In L. Bickman (Ed.), *Using program theory in evaluation*. San Francisco: Jossey-Bass.

Bickman, L. (2000). Summing up program theory. *New Directions for Evaluation, 2000*(87), 103-112.

Birkmayer, J. D., & Weiss, C. H. (2000). Theory-Based Evaluation in Practice: What do we learn? *Evaluation Review, 24*, 407-431.

Boonekamp, G., Colomer, C., Tomas, A., & Nunez, A. (1999). Healthy Cities Evaluation: the co-ordinators perspective. *Health Promotion International, 14*(2), 103-187.

Boutilier, M., Cleverly, S., & Labonte, R. (2000). Community as a Setting for Health Promotion. In B. D. Poland, L. W. Green & I. Rootman (Eds.), *Settings for Health Promotion: Linking theory and practice*. Thousand Oaks: Sage.

Braveman, P. A., Egerter, S. A., Woolf, S. H., & Marks, J. S. (2011). When Do We Know Enough to Recommend Action on the Social Determinants of Health? *American Journal of Preventive Medicine, 40*(1S1), S58-S66.

Bryman, A. (2012). *Social Research Methods* (4th ed.). New York: Oxford University Press.

Burton, S. (1999). Evaluation of Healthy City projects: stakeholder analysis of two projects in Bangladesh. *Environment and Urbanization, 11*(1), 41-52.

Byng, R., Norman, I., & Redfern, S. (2005). Using Realistic Evaluation to Evaluate a Practice-level Intervention to Improve Primary Healthcare for Patients with Long-term Mental Illness. *Evaluation, 11*(1), 69-93.

Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., . . . Kinmonth, A. (2007). Designing and evaluating complex interventions to improve health care. *British Medical Journal, 334*(7591), 455-459.

Catford, J. (2011). Ottawa 1986: back to the future. *Health Promotion International, 26*(suppl 2), ii163-ii167.

Chen, H.-T. (1990). *Theory-driven Evaluations*. Newbury Park, California: Sage Publications Inc.

Chen, H.-T. (2005). *Practical Program Evaluation: assessing and improving planning, implementation and effectiveness*. Thousand Oaks, California: Sage Publications Inc.

Consumer Focus Collaboration. (2000). *Feedback, Participation and Consumer Diversity: a literature review*. Canberra: Commonwealth Department of Health and Aged Care.

Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation, 32*(2), 199-226.

Crotty, M. (1998). *The Foundations of Social Research: meaning and perspective in the research process*. St Leonards, NSW: Allen and Unwin.

de Leeuw, E. (1999). Monitoring, accountability, reporting, impact assessement: a framework for Phase 111 Healthy Cities: Centre for Urban Health, WHO, Regional Office for Europe.

de Leeuw, E., & Skovgaard, T. (2005). Utility-driven evidence for Healthy Cities: Problems with evidence generation and application. *Social Science & Medicine, 61*(6), 1331-1341.

Department of Human Services. (2001). Metropolitan Community Health Services Review: Final Report. Adelaide: Department of Human Services, Metropolitan Division.

Donaldson, S. (2003). Theory-driven program evaluation in the new millennium. In S. Donaldson & M. Scriven (Eds.), *Evaluating Social Programs and Problems: Visions for the new millennium*. New Jersey: Lawrence Eribaum Associates.

Dooris, M. (1999). Healthy Cities and Local Agenda 21: the UK experinece - challenges for the new Millennium. *Health Promotion International, 14*(4), 365-375.

Dooris, M. (2005). Healthy Settings: challenges to generating evidence of effectiveness. *Health Promotion International, 21*(1), 55-65.

Dooris, M. (2012). The settings approach: looking back, looking forward. In A. Scriven & M. Hodgins (Eds.), *Health Promotion Settings: Principles and Practice*. London: Sage.
Evans, L., Hall, M., Jones, C. M., & Neiman, A. (2007). Did the Ottawa Charter play a role in the push to assess the effectiveness of health promotion? *Promotion and Education, Supplement*(2), 28-30.
Fishman, D. (1992). Postmodernism comes to program evaluation: A critical review of Guba and Lincoln's *Fourth Generation Evaluation*. *Evaluation and Program Planning, 15*(3), 263-270.
Fry, D., & Baum, F. (1992). Keywords in Community Health. In F. Baum, D. Fry & I. Lennie (Eds.), *Community Health Policy and Practice in Australia*. Leichardt, NSW: Pluto Press Australia.
Funnell, S., & Rogers, P. J. (2011). *Purposeful Program Theory: Effective use of theories of change and logic models*. San Fransisco: Jossey-Bass.
Furler, E. (1979). Against hegemony in health care service evaluations. *Community Health Studies, 3*(1), 32-41.
Government of South Australia. (2006). Healthy Ageing - Nutrition: Project Report. Adelaide.
Green, G., & Tones, K. (2010). *Health Promotion: Planning and Strategies* (2nd ed.). London: Sage.
Green, L. W. If we want more evidence-based practice, we need more practice-based evidence. Published Applications Of The Precede Model Retrieved 18 June 2012, from http://lgreen.net/precede%20apps/preapps-NEW.htm
Green, L. W., & Kreuter, M. W. (1999). *Health Promotion Planning: an educational and ecological perspective* (3rd ed.). California: Mayfield Publishing Company.
Green, L. W., Poland, B. D., & Rootman, I. (2000). The Settings Approach to Health Promotion. In B. D. Poland, L. W. Green & I. Rootman (Eds.), *Settings for Health Promotion: Linking theory and practice*. Thousand Oaks, California: Sage.
Greenhalgh, T., Humphrey, C., Hughes, J., MacFarlane, F., Butler, C., & Pawson, R. (2009). How Do You Modernize a Health Service? A Realist Evaluation on Whole-Scale Tranformation in London. *The Millbank Quarterly, 87*(2), 391-416.
Gregory, A. (2000). Problematizing Participation. *Evaluation, 6*(2), 179-199.
Grembowski, D. (2001). *The Practice of Health Program Evaluation*. Thousand Oaks, California: Sage Publicactions, Inc.
Guba, E., & Lincoln, Y. (1989). *Fourth Generation Evaluation*. Newbury Park, California: Sage Publications, Inc.
Hancock, T. (1986). Lalonde and beyond: Looking back at “A New Perspective on the Health of Canadians”. *Health Promotion International, 1*(1), 93-100.
Hancock, T. (1993). The Healthy City from concept to application: Implications for research. In J. K. Davies & M. P. Kelly (Eds.), *Healthy Cities: Research and Practice*. London: Routledge.
Hancock, T. (2011). Health promotion in Canada: 25 years of unfulfilled promise. *Health Promotion International, 26*(suppl 2), ii263-ii267.
Hawe, P., Bond, L., & Butler, H. (2009). Knowledge theories can inform evaluation practice: What can a complexity lens add? *New Directions for Evaluation*. (124), 89-100.

Hawe, P., Degeling, D., & Hall, J. (1990). *Evaluating Health Promotion: a health worker's guide*. Artarmon, NSW: MacLennan and Petty Pty Ltd.

Hawe, P., Noort, M., King, L., & C, J. (1997). Multiplying Health Gains: the critical role of capacity-building within health promotion programs. *Health Policy, 39*(1), 29-42.

Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how 'out of control' can a randomised controlled trial be? *British Medical Journal, 328*, 1561-1563.

Hawe, P., Shiell, A., & Riley, T. (2009). Theorising interventions as events in systems. *American Journal of Community Psychology, 43*, 267-276.

Howat, P., Jones, S., Hall, M., Cross, D., & Stevenson, M. (1997). The PRECEDE-PROCEED model: application to planning a child pedestrian injury prevention program. *Injury Prevention, 3*(4), 282-287.

International Union for Health Promotion and Education. (2000). The evidence of health promotion effectiveness: shaping public health in a new Europe. A report for the European Commission / by the International Union for Health Promotion and Education. Part Two: Evidence Book. Health Promotion Effectiveness: The questions to be asked. Brussels: European Commission.

Israel, B. A., Schultz, A. J., Parker, E. A., & Becker, A. B. (1998). Review of Community-Based Research: Assessing partnership approaches to improve public health. *Annual Review of Public Health, 19*, 173-202.

Jewkes, R., & Murcott, A. (1998). Community representatives: Representing the "community"? *Social Science and Medicine, 46*(7), 843-858.

Jirojwong, S., & Liamputtong, P. (2009a). Population Health and Health Promotion. In S. Jirojwong & P. Liamputtong (Eds.), *Population Health, Communities and Health Promotion*. Melbourne: Oxford University Press.

Jirojwong, S., & Liamputtong, P. (2009b). Primary Health Care and Health Promotion. In S. Jirojwong & P. Liamputtong (Eds.), *Population Health, Communities and Health Promotion*. Melbourne: Oxford University Press.

Johnson, A., & Baum, F. (2001). Health promoting hospitals: a typology of different organizational approaches to health promotion. *Health Promotion International, 16*(3), 281-287.

Jolley, G. (2003). *If only numbers count: the use of performance indicators as an evaluation tool for community health servcies.* M.Sc.(PHC) Masters Thesis, Flinders University, Adelaide.

Jolley, G., Baum, F., Hurley, C., & Fry, D. (2004). What Works? A systematic review of the effectiveness of community health in Adelaide. Adelaide: South Australian Community Health Research Unit.

Jolley, G., Lawless, A., Hurley, C., Biedrzycki, K., & Ramanathan, R. (2007). *Evaluation Toolkits for Equity, Community Participation and Collaborative Partnerships* Paper presented at the Australian Health Promotion Association 17th National Conference, 1 - 4 May 2007, Adelaide.

Jones, S. C., & Donovan, R. J. (2004). Does theory inform practice in health promotion in Australia? *Health Education Research, 19*(1), 1-14.

Judd, J., Frankish, C. J., & Moulton, G. (2001). Setting standards in the evaluation of community-based health promotion programmes— a unifying approach. *Health Promotion International, 16*(4), 367-380.

Judge, K., & Bauld, L. (2001). Strong theory, flexible methods: evaluating complex community-based initiatives. *Critical Public Health, 11*(1), 19-38.

Judge, K., & Mackenzie, M. (2002). Theory-based evaluation: New approaches to evaluating complex community-based initiatives. In J. Mackenbach & M. Bakker (Eds.), *Reducing Inequalties in Health: A European perspective*. London: Routledge.

Kavanagh, A., Daly, J., & Jolley, D. (2002). Research methods, evidence and public health. *Australian and New Zealand Journal of Public Health, 26*(4), 337-342.

Keleher, H. (2007). Health Promotion Planning and the Social Determinants of Health. In H. Keleher, C. MacDougall & B. Murphy (Eds.), *Understanding Health Promotion*. Melbourne: Oxford University Press.

Keleher, H., MacDougall, C., & Murphy, B. (2007). Approaching Health Promotion. In H. Keleher, C. MacDougall & B. Murphy (Eds.), *Understanding Health Promotion*: Oxford University Press.

Kelly, M. P., Morgan, A., Bonnefoy, J., Butt, J., & Bergman, V. (2007). The Social Determinants of Heatlh: developing an evidence base for political action. Final Report to World Health Organization Commission on the Social Determinants of Health. Measurement and Knowledge Network: WHO.

Kemmis, S., & McTaggart, R. (1988). *The Action Research Planner*. Melbourne: Deakin University.

Keshavarz, N., Nutbeam, D., Rowling, L., & Khavarpour, F. (2010). Schools as social complex adaptive systems: A new way to understand the challenges of introducing the health promoting schools concept. *Social Science and Medicine, 70*(10), 1467-1474.

Kickbusch, I. (1997). Think health: What Makes the Difference? *Health Promotion International, 12*(4), 265-272.

Kökény, M. (2011). Ottawa revisited: 'enable, mediate and advocate'. *Health Promotion International, 26*(suppl 2), ii180-ii182.

Kremser, W. (2011). Phases of school health promotion implementation through the lens of complexity theory: lessons learnt from an Austrian case study. *Health Promotion International, 26*(2), 136-147.

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal, 42*(3), 462-483.

Labonte, R. (1992). Heart Health Inequalities in Canada: Models, theories and planning. *Health Promotion International, 7*(2), 119-127.

Labonte, R., & Penfold, S. (1981). Canadian perspectives in health promotion: A critique. *Health Education 19*(3/4), 4-9.

Lalonde, M. (1974). A New Perspective on the Health of Canadians. Ottawa: Minister of Supply and Services, Canada.

Laris, P., Jolley, G., & Smith, A. (2001). Eat Well SA Evaluation Report October 1997-October 2000. Adelaide.

Laughlin, R., & Broadbent, J. (1996). Redesigning Fourth Generation Evaluation. *Evaluation, 2*(4), 431-451.

Ling, T. (2012). Evaluating complex and unfolding interventions in real time. *Evaluation, 18*(1), 79-91.

Mackenzie, M., & Blamey, A. (2005). The Practice and the Theory: Lessons from the application of a theories of change approach. *Evaluation, 11*(2), 151-168.

Mark, M. (2003). Toward an Integrative View of the Theory and Practice of Program and Policy Evaluation. In S. Donaldson & M. Scriven (Eds.), *Evaluating Social Programs and Problems: Visions for the new millennium*. New Jersey: Lawrence Eribaum Associates.

Martin, C. M., & Félix-Bortolotti, M. (2010). W(h)ither complexity? The emperor's new toolkit? Or elucidating the evolution of health systems knowledge? *Journal of Evaluation in Clinical Practice, 16*(3), 415-420.

Matheson, A., Dew, K., & Cumming, J. (2009). Complexity, evaluation and the effectiveness of community-based interventions to reduce health inequalities. *Health Promotion Journal of Australia, 20*(3), 221-221-226.

McLeroy, K., Bibeau, D., Steckler, A., & Glanz, K. (1988). An Ecological Perspective on Health Promotion Programs. *Health Education Quarterly, 15*(4), 351-377.

McQueen, D. V. (2007). Evidence and Theory: Continuing debates on evidence and effectiveness. In D. V. McQueen & C. M. Jones (Eds.), *Global Perspectives on Health Promotion Effectiveness*. New York: Springer.

Medical Research Council. (2008). Developing and evaluating complex interventions: new guidance. In MRC (Series Ed.) Retrieved from www.mrc.ac.uk/complexinterventionsguidance

Metzler, M., Amuyunzu-Nyamongo, M., Mukhopadhyay, A., & De Salazar, L. (2007). Community interventions on social determinants of health: focusing the evidence In D. V. McQueen & C. M. Jones (Eds.), *Global perspectives on health promotion effectiveness* New York: Springer.

Minkler, M., Glover Blackwell, A., Thompson, M., & Tamir, H. (2003). Community-based participatory research: Implications for public health funding. *American Journal of Public Health, 93*(8), 1210-1213.

Mullen, P. D., Evans, D., Forster, J., Gottlieb, N. H., Kreuter, M., Moon, R., . . . Strecher, V. J. (1994). Settings as an Important Dimension in Health Education/Promotion Policy, Programs, and Research. *Health Education & Behavior, 22*(3), 329-345.

Neiman, A., & Hall, M. (2007). Urbanization and Health Promotion: Challenges and opportunities. In D. V. McQueen & C. M. Jones (Eds.), *Global Perspectives on Health Promotion Effectiveness*. New York: Springer.

Norman, C. D. (2009). Health promotion as a systems science and practice. *Journal of Evaluation in Clinical Practice, 15*(5), 868-872.

Nunn, R. J. (2007). Complexity theory applied to itself. *Emergence: Complexity and Organization, 9*(1/2), 93-106.

Nutbeam, D. (1996). Health Outcomes and Health Promotion - Defining Success in Health Promotion. *Health Promotion Journal of Australia, 6*(2), 58-60.

Nutbeam, D. (1998). Evaluating Health Promotion—Progress, Problems and Solutions. *Health Promotion International, 13*(1), 27-44.

Nutbeam, D. (1999). The challenge to provide 'evidence' in health promotion. *Health Promotion International, 14*(2), 99-101.

Nutbeam, D., Harris, E., & Wise, M. (2010). *Theory in a Nutshell: a practical guide to health promotion theories* (3rd ed.). North Ryde, NSW: McGraw-Hill Australia Pty Ltd.

Oakley, A. (1998). Experimentation and social interventions: a forgotten but important history. *BMJ, 317*(7167), 1239-1242.

Oakley, A. (2005). Design and Analysis of Social Intervention Studies in Health Research. In A. Bowling & S. Ebrahim (Eds.), *Handbook of Health Research Methods: Investigation, Measurement and Analysis*. Berkshire: OUP.

Oakley, P., & Kahassay, H. (1999). Community Involvement in Health Development: an overview. In H. Kahassay & P. Oakley (Eds.), *Community*

Involvement in Health Development: a review of the concept and practice. Geneva: World Health Organization.

Owen, A., & Lennie, I. (1992). Health for all and community health. In F. Baum, D. Fry & I. Lennie (Eds.), *Community Health: Policy and Practice in Australia*. Sydney: Pluto Press.

Patton, M. (1997). *Utilization-Focused Evaluation: the new century text* (3rd ed.). Thousand Oaks, California: Sage.

Patton, M. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Thousand Oaks, California: Sage.

Patton, M. (2011). *Developmental Evaluation: Applying compexity concepts to enhance innovation and use*. New York: The Guildford Press.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

Peersman, G. (2001). Promoting Health: principles of practice and evaluation. In S. Oliver & G. Peersman (Eds.), *Using Research for Effective Health Promotion*. Buckingham: Open University Press.

Petersen, A., & Lupton, D. (1996). *The New Public Health: health and self in the age of risk*. St Leonards, NSW: Allen and Unwin.

Pettersson, B. (2011). Some bitter-sweet reflections on the Ottawa Charter commemoration cake: a personal discourse from an Ottawa rocker. *Health Promotion International, 26*(suppl 2), ii173-ii179.

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *British Medical Journal, 323*, 625-628.

Pluye, P., Potvin, L., Denis, J. L., & Pelletier, J. (2004). Program sustainability: focus on organizational routines. *Health Promotion International, 19*(4), 489-500.

Poland, B. D. (1996a). Knowledge development and evaluation in, of and for Healthy Community initiatives. Part I: guiding principles. *Health Promotion International, 11*(3), 237-247.

Poland, B. D. (1996b). Knowledge development and evaluation in, of and for Healthy Community initiatives. Part II: potential content foci. *Health Promotion International, 11*(4), 341-349.

Poland, B. D., Frohlich, K. L., & Cargo, M. (2009). Context as a Fundemental Dimension of Health Promotion Program Evaluation. In L. Potvin, D. V. McQueen, M. Hall, L. de Salazar, L. Anderson & Z. Hartz (Eds.), *Health Promotion Practices in the Americas: Values and Research*. New York: Springer.

Poland, B. D., Krupa, G., & McCall, D. (2009). Settings for Health Promotion: An Analytic Framework to Guide Intervention Design and Implementation. *Health Promotion Practice, 10*(4), 505-516.

Popay, J., Rogers, A., & Williams, G. (1998). Rationale and standards for sytematic review of qualitative literature in health services research. *Qualitative Health Research, 8*, 341-351.

Potvin, L., & Bisset, S. (2009). There is More to Methodology than Method. In L. Potvin, D. V. McQueen, M. Hall, L. de Salazar, L. Anderson & Z. Hartz (Eds.), *Health Promotion Practices in the Americas: Values and Research*. New York: Springer.

Potvin, L., & McQueen, D. V. (2009). Dilemmas for Health Promotion Evaluation. In L. Potvin, D. V. McQueen, M. Hall, L. de Salazar, L. Anderson & Z. Hartz (Eds.), *Health Promotion Practices in the Americas: Values and Research*. New York: Springer.

Public Health Research Advisory Committee. (2008). Report of the Review of Public Health Research Funding in Australia. Canberra: National Health & Medical Research Council.

Raftery, J. (1995). The social and historical context. In F. Baum (Ed.), *Health for All: The South Australian Experience*. Adelaide: Wakefield Press.

Reason, P. (1994). Human Inquiry as Discipline and Practice. In P. Reason (Ed.), *Participation in Human Inquiry: Research with People*. London: Sage.

Rifkin, S. (1996). Paradigms Lost: Toward a new understanding of community participation in health programmes. *Acta Tropica, 61*(2), 79-92.

Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsi, T. A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation, 2000*(87), 5-13.

Rootman, I., Goodstadt, M., Potvin, L., & Springett, J. (1997). Towards a Framework for Health Promotion Evaluation. Copenhagen: World Health Organization, Regional Office for Europe.

Rootman, I., Goodstadt, M., Potvin, L., & Springett, J. (2001). A framework for health promotion evaluation. In I. Rootman, M. Goodstadt, B. Hyndman, D. V. McQueen, L. Potvin, J. Springett & E. Ziglio (Eds.), *Evaluation in Health Promotion: Principles and Perspectives*. Copenhagen: World Heath Organization.

Rossi, P., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). California: Sage Publications Inc.

Rychetnik, L., & Frommer, M. (2001). A proposed schema for evaluating evidence on public health interventions. Version 3.1. Melbourne: National Public Health Partnership.

Saan, H., & Wise, M. (2011). Enable, mediate, advocate. *Health Promotion International, 26*(suppl 2), ii187-ii193.

Scriven, M. (1981). *Evaluation Thesaurus* (3rd ed.). California: Edgepress.

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

Sharp, C. (2003). Development of program evaluation in Australia and the Australasian Evaluation Society - the early decades. *Evaluation Journal of Australasia, 3*(2), 6-16.

Shediac-Rizkallah, M. C., & Bone, L. R. (1998). Planning for the sustainability of community-based health programs: conceptual frameworks and future directions for research, practice and policy. *Health Education Research, 13*(1), 87-108.

Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. *BMJ, 336*(7656), 1281-1283.

Signal, L. N., Walton, M. D., Ni Mhurchu, C., Maddison, R., Bowers, S. G., Carter, K. N., . . . Pearce, J. (2012). Tackling 'wicked' health promotion problems: a New Zealand case study. *Health Promotion International*.

Smithies, J., & Adams, L. (1993). Walking the Tightrope: issues in evaluation and community participation for Health for All. In J. K. Davies & M. P. Kelly (Eds.), *Healthy Cities: Research and Practice*. London: Routledge.

South Australian Community Health Research Unit. (2000). *Improving Health Services through Consumer Participation: a resource guide for organisations*. Canberra: Commonwealth Department of Health and Aged Care.

South Australian Community Health Research Unit. (2012). South Australian Community Health Research Unit Retrieved 25/3/2012, from http://www.flinders.edu.au/medicine/sites/sachru/

South Australian Community Health Research Unit. (undated). Are you writing an evaluation report? Evaluation Report Template. Adelaide: South Australian Community Health Research Unit, Flinders University.

South, J., & Woodall, J. (2012). Planning and evaluating health promotion in settings. In A. Scriven & M. Hodgins (Eds.), *Health Promotion Settings: Principles and Practice*. London: Sage.

Talbot, L., & Verrinder, G. (2010). *Promoting Health: the primary health care approach* (4th ed.). Sydney: Churchill Livingstone.

Tones, K., & Green, J. (2004). *Health Promotion: Planning and strategies*. London: Sage.

Tremblay, M.-C., & Richard, L. (2011). Complexity: a potential paradigm for a health promotion discipline. *Health Promotion International*.

Walsh, J. A., & Warren, K. S. (1979). Selective Primary Health Care. *New England Journal of Medicine, 301*(18), 967-974.

Wass, A. (2000). *Promoting Health: the primary health care approach* (2nd ed.). Sydney: Saunders.

Weiss, C. (1998). *Evaluation* (2nd ed.). New Jersey: Prentice Hall.

Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation, 18*(4), 405-420.

Westley, F., Zimmerman, B., & Patton, M. (2006). *Getting to Maybe: How the world is changed*: Random House Canada.

Whitehead, M. (1991). The concepts and principles of equity and health. *Health Promotion International, 6*(3), 217-228.

Wilkinson, R. G. (1996). *Unhealthy Societies. The Afflictions of Inequality*. London: Routledge.

Wimbush, E., & Watson, J. (2000). An Evaluation Framework for Health Promotion: Theory, Quality and Effectiveness. *Evaluation, 6*(3), 301-321.

World Health Organization. (1986a). Ottawa Charter for Health Promotion. *Health Promotion International, 1*(4), i-v.

World Health Organization. (1986b). Social justice and equity in health: Report on WHO meeting, Copenhagen DNK Regional Office for Europe, World Health Organization.

World Health Organization. (1997). *Jakarta Declaration on Leading Health Promotion into the 21st Century.* Paper presented at the The Fourth International Conference on Health Promotion: New Players for a New Era - Leading Health Promotion into the 21st Century, 21 to 25 July 1997 Jakarta

World Health Organization. (1998). Health Promotion Glossary. Geneva: World Health Organization.

World Health Organization. (1999). *Preparing for the 21st Century*. Malacca, Malaysia.

World Health Organization. (2000). Healthy Cities in Action: 5 case studies from Africa, Asia, Middle East and Latin America. Geneva: Department of the Human Environment, WHO.

World Health Organization. (2005). The Bangkok Charter for Health Promotion in a Globalized World. Retrieved from doi:http://www.who.int/healthpromotion/conferences/6gchp/hpr_050829_%20BCHP.pdf

World Health Organization. (2008). The world health report 2008 : primary health care now more than ever: introduction and overview Geneva.
Wright, A., McGorry, P., Harris, M., Form, A., & Pennell, K. (2006). Development and evaluation of a youth mental health community awareness campaign - The Compass Strategy. *BMC Public Health, 6*, 215-227.
Young, I. (1990). The ideal of community and the politics of difference. In J. J. Nicholson (Ed.), *Feminism/Postmodernism*. New York: Routledge.

APPENDIX: PUBLICATIONS FORMING PART OF THE THESIS

1. Smith, A., Coveney, J., Carter, P., Jolley, G. and Laris, P. (2004) The Eat Well SA project: an evaluation-based case study in building capacity for promoting healthy eating. *Health Promotion International* 19 (3) 327-334.

2. Baum, F. Jolley, G. Hicks, R. Saint, K. & Parker, S. (2006) What makes for sustainable Healthy Cities initiatives? - a review of the evidence from Noarlunga after 18 years, *Health Promotion International*, 21 (4) 259-265.

3. Jolley, G. Lawless, A. Baum, F. Hurley, C. and Fry, D. (2007) Building an evidence base for community health: a review of the quality of program evaluations. *Australian Health Review*, 31 (4) 603-610.

4. Jolley G. Lawless A and Hurly C. (2008) Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion. *Health Promotion Journal of Australia*, 19 (2) 152-157.

5. Jolley G. (2008) Evaluation of an action research project in workforce development and organisational change: Healthy Ageing-Nutrition *Evaluation Journal of Australasia*, 8 (1) 11-19.