CHAPTER 1
INTRODUCTION TO THE STUDY

> There are two key reasons for evaluating the performance of schools in any school system. The first is **accountability**—whether that accountability is to the Government, the taxpayer, or to the students themselves. The underlying principle is that schools, being in the main publicly funded and having a role to play in the education of present and future generations, should fulfil the aims, purposes and objectives which society defines for them.
>
> The second, related reason is **school improvement**. A school's performance needs to be evaluated or assessed in order to identify current strengths and areas for development as a basis for identifying and implementing an appropriate strategy for improvement. If support to schools, including training and development, is to be provided where it is most needed, it is necessary to identify needs and priorities.
>
> (New South Wales Department of School Education, n.d.-b:1)

THE STUDY

Evaluation can serve both summative and formative purposes. To these ends a pilot program of Cyclical Reviews was introduced into government schools in the Western Sydney Region of New South Wales during the period 2006–09. The Cyclical Review process was developed and underwent continual refinement through action research. This study is about the influence these pilot Cyclical Reviews had on the school principals who participated in them.

BACKGROUND TO THE STUDY

School Reviews

Evaluating the performance of schools is not new—in neither New South Wales, nor Australia, nor other countries. The structures, however, for evaluating schools and the balance in these structures between accountability and improvement have, in recent years, been the subject of much research and debate. The context of these ideas is discussed in Chapter 2.

One component, the school review, regularly featured in these structures, although in a range of different forms. In New South Wales government schools, as is the case in schools across Australia, school reviews, which form one component of broader school accountability and development approaches, are evolving.

This evolution is in part due to the growing expectation that schools develop an organisational culture in which data-driven planning and robust evaluation are the norm. In New South Wales government schools, school leaders are required to analyse student-performance data, undertake evaluations of key curriculum and management areas, and set improvement targets (New South Wales Department of Education and Training, 2005). To assist in this task, ranges of data, including sophisticated electronic data-analysis tools, are provided to schools, on-line resources are available on a range of school-development and self-evaluation topics, and in some cases regional officers with skills in data analysis or evaluation provide direct support to schools.

Although there is a system-developed and regionally implemented process contained in the *School Development Policy* (New South Wales Department of Education and Training, 2004) to review, on the basis of available data, underperforming schools or programs within schools, there has been no system-supported mechanism for all

schools to participate in, or be subject to, regular reviews of their whole-school operation or to develop their own evaluation capacity. Such a process was foreshadowed in 2003–04 with the production of a number of in-house and unpublished papers regarding the future structure of school accountability. These resulted in the publication of a new system-wide framework on the New South Wales Department of Education and Training's Intranet—*Framework for School Development and Accountability* (New South Wales Department of Education and Training, 2005)—where the component called 'cyclic reviews' appeared, but no definition or specification of what a cyclic review entailed was given, and no further work on these cyclic reviews at a system-wide level was pursued at that time.

Cyclical Reviews

At a regional level, however, the Director of Western Sydney Region (one of ten school regions within the New South Wales Department of Education and Training) took the initiative in 2006 to develop the cyclic review component of the framework for that region and listed the development of what he termed a regional 'Cyclical Review' process for schools that would 'strengthen school development support and community confidence' as a priority in its 2006–08 strategic plan (Western Sydney Region New South Wales Department of Education and Training, 2005).

From 2006 to 2009 this process was developed and piloted. Implicit in the stated purpose of Cyclical Reviews were multiple summative and formative purposes. Cyclical Reviews were to develop a mechanism for school evaluation that would not only give a robust, useful, and influential evaluation of whole-school performance and governance for the school being reviewed, but also provide a sustainable regional framework in terms of time, personnel, and cost for such reviews, build the evaluation capacity of principals, and school and regional support staff, and strengthen community confidence in regional government schools (Wasson, 2006a, meeting 1 March).

The first three of the above-mentioned purposes, namely to develop a mechanism for

- a useful and influential evaluation of whole-school performances and governance
- a sustainable framework for such reviews
- building the evaluation capacity of principals,

are the key foci of this study. Each purpose will be examined through the involvement of principals in the process, as it was principals who were considered to be pivotal in achieving these purposes. This consideration arose from the increased responsibility for accountability, evaluation, and school effectiveness and improvement that had been imposed on New South Wales government-school principals through the Crown Employees Salaries and Conditions Award of 2004 (an industrial agreement between the Department and the New South Wales Teachers Federation that sets the salary for principals subject to their meeting specific conditions).

In order to examine the involvement of principals in the evaluation processes, the concept of evaluation influence was seen as particularly relevant. Given that school improvement was seen as a major reason for the introduction of Cyclical Reviews, the use that the principals made of the evaluation findings and whether they were influenced by these findings was important. A further matter, of equal importance, was the influence that involvement in the review process had on the principals' skills in, understandings of, and attitudes towards evaluation. This influence was interpreted as being not only that experienced by the principals as individuals but also the influence they in turn exerted on others. Finally, if the Cyclical Reviews could be seen as enhancing the skills, understandings, attitudes, and application of evaluation processes and procedures by those involved, then that enhancement is defined as evaluation capacity building (ECB).

PURPOSE OF THE STUDY

This study has four specific purposes.

1. To identify and interpret how the eighteen principals who had participated in the pilot Cyclical Review program in Western Sydney Region had been influenced over time by both the review process and review findings.
2. To gain insights into how, when, and to what extent Cyclical Reviews shape, effect, support, and change the views and practices of the principals who participate in them.
3. To use the knowledge gained from 1 and 2 above to modify and strengthen the Cyclical Review program in Western Sydney Region and, in so doing, contribute to the statewide implementation of that element of the New South Wales *School Development and Accountability Framework*.
4. To add to current knowledge concerning Cyclical Reviews.

SIGNIFICANCE OF THE STUDY

In recent years there has been an emergence of innovative approaches to evaluation. Such approaches have heralded a shift in focus to more participatory, qualitative research that offers a lens through which to interpret the results of programs and to provide valuable knowledge to guide program improvement (Delaney, 2006). This shift is occurring against the backdrop of a discussion around the need to examine the effect of evaluation by expanding the notion of evaluation use to a broader notion of evaluation influence (Cousins & Shula, 2006; Henry & Mark, 2003; Kirkhart, 2000; Mark and Henry, 2004). This research further tests the concept of influence in relation to Cyclical Reviews.

The research also focuses on developing a better understanding of the factors contributing to longer-term evaluation influence and the relationships between these seemingly interdependent factors.

The researcher has worked in the now New South Wales Department of Education and Communities for over thirty-five years and since 2001 has been working in the area of school improvement, accountability, and development. A key feature of her work has been the leading of reviews under the *School Development Policy* (New South Wales Department of Education and Training, 1999, 2004) and in 2003 she managed, as it existed at that time, the statewide program of reviews. From this work the researcher developed a belief that a school review, conducted under the *School Development Policy*, could be a form of evaluation that assisted schools to improve. Anecdotal evidence, however, suggested that the initial flurry of activity that had followed the reviews conducted under this policy dissipated rather quickly—within about six months, in fact. Again anecdotally, this often appeared to happen because the conditions that were in place when the review was undertaken changed (for example, school staff at the time of the review retired or moved elsewhere) and the long-term effect—that is, beyond about six months—appeared to be minimal. The key problem to be researched was therefore the impact of Cyclical Reviews.

A cost–benefit analysis indicated the significance of the problem. Reviews conducted within the *School Development Policy* are costly in terms of both human and physical resources. Each review requires a senior Departmental school improvement officer to prepare for, lead, and write the report for the review—in total approximately three weeks' work. Further, each review team comprises a minimum of four other school staff, each requiring relief to be paid to their school during the training and review period—approximately one week per team member. In addition, depending on the location, a variety of physical resources—computer hardware and software, print-based materials, travel, accommodation, and catering, are required. Despite this, to date there has been no formal evaluation of the use made of school reviews or the

influence of individual reviews on individual participants, the school itself, or on the system.

It is anticipated that the contribution of this research to knowledge will therefore be twofold: it will further contribute to the understanding of evaluation influence; and in particular, it will provide new knowledge about the influence of the Cyclical Reviews of Western Sydney Region on participating principals and how this influence could be improved.

SCOPE OF THE STUDY

In 2006, in line with the Western Sydney Region strategic plan, work commenced on developing the Cyclical Review process and on gaining the commitment of principals. In 2007 a Cyclical Review working-party comprising nine school principals (who had been involved in the development during 2006) and the researcher, in her then role of regional school development officer, scheduled a series of pilot Cyclical Reviews to take place in 2007 and early 2008. Each of the nine principals agreed to his/her school's undergoing a Cyclical Review and that each of them would lead a Cyclical Review in one other school in the pilot scheme. In accordance with the agreed procedures for the conduct of these reviews a further nine principals from across the region were also co-opted to be team members of one of the reviews. The study set out to collect data from each of these eighteen principals concerning their involvement from 2006 until mid-2009. In terms of understanding the timelines associated with this research it needs to be noted that the researcher needed to defer final analysis and write up owing to unforeseeable family circumstances in the period from late 2010 to mid-2013.

Definition of Key Terms

Department

The New South Wales Government department responsible for school education has had a series of name changes since the introduction of compulsory school education in New South Wales in 1848, including the Department of Education, the New South Wales Department of School Education, the New South Wales Department of Education and Training, and the New South Wales Department of Education and Communities as it is currently known. For the purposes of this study the term *Department* has been adopted to refer to the government department responsible for school education in New South Wales by whatever name it was known at the time (including the time when the Cyclical Review process took place), unless the specific name needs to be retained for the historical context. All other government departments and ministries nationally and internationally are referred to in full.

Evaluation influence

The terms e*valuation utilisation* and *use* have been adopted in the literature (as discussed in Chapter 3) since the 1970s to describe the diverse effects of evaluation. With an initial focus on instrumental use, the range of meanings of these terms first expanded to include conceptual and symbolic use of evaluation results and later to include the use made of the evaluation process. More recently the concept of *evaluation influence* has been proposed. Kirkhart (2000) proposed this term to expand the concept of use so that it included non-results-based impact, unintended results, and incremental impact. Mark and Henry (2004) adopted this term because they believed that it included changes that occur at the location and time of the evaluation itself, changes that take place elsewhere and later, and changes that occur at the individual, interpersonal, and collective level.

While the literature shows apparent differences, both conceptually and linguistically, between use and influence, there is as yet no agreement on which term should take precedence or in which contexts the different terms should be used. In the context of this study, the term *evaluation influence* has been adopted to refer to the influences exerted on the principals in the preparation for, during, and after their participation in Cyclical Reviews as either intended or unintended consequences on them as individuals, as leaders in their schools, or as senior officers in the region.

Cyclical Review

School reviews in various forms have long been a component of school accountability and development frameworks. At different stages of public education in New South Wales, reviews have been variously termed inspections, Quality Assurance reviews, and reviews under the *School Development Policy*, where they are termed specifically *school education support team visits*, *program reviews*, and *management reviews*. These latter three reviews are all reviews by exception, occurring only when a specified need arises; for example, when there is dysfunction in the management of the school. As indicated above the Department's accountability framework of 2005 foreshadowed the cyclic review of all schools. In 2006–07 the Western Sydney Region of the Department developed its own response to this component of the framework and adopted the term *Cyclical Review*, and this term is therefore used exclusively in this study to refer to the process adopted in Western Sydney Region. The process developed and the reviews conducted under this process in the period 2006–09 and which are the subject of this study constituted a pilot scheme. For simplicity in this study the term *pilot* will be dropped when referring to these reviews.

Participating principals

The term *participating principals* refers to the principals involved in the planning or implementation or both of the Cyclical Review program between 2006 and 2009.

External and internal reviews and reviewers

The terms *external review*, *internal review*, *external reviewers*, and *internal reviewers* are used with different meanings by various researchers and systems. Thus, for example, a review conducted by personnel working for an organisation but not directly working in the functional area or department under review has been variously referred to as an external or an internal review using external or internal reviewers. A third term, *independent*, is also used in Australian settings. Cuttance notes that

> many Australian systems use the term 'external' to refer to staff who are external to the school (e.g. District Director), as distinct from 'independent', that is professionals with no stake or direct self-interest in the performance of schools and outcomes of the accountability system.
>
> (Radii, 2005:5)

This distinction is a useful one. Cyclical Review teams comprise both those internal to and external to the school, all of whom have a direct interest in school performance school development, and the accountability system. Using the distinction drawn by Cuttance, the team-leader principals and team-member principals in this study are external reviewers, while the host principal is an internal reviewer. The review team comprises more external than internal members. As the review process is meant to apply a 'fresh set of eyes' to the school's operations, and as the report is a public one, and as the recommendations require action and public reporting (in the form of a section in the school's annual report), the review itself is defined, for the purposes of this study, as a type of external review.

PPODS

Pocket PCs for Organising Data and Sorting (PPODS) is a descriptive term given to the method developed for the Cyclical Review process to collect and record information from interviews, observations, and documents. Using PPODS, the recorder writes directly onto the screen of the pocket PC. Each discrete piece of information is coded and separated by carriage returns, allowing the data to be transferred to a computer spreadsheet for sorting and analysis. This same method was used by the researcher to collect, sort, and analyse the data for this study.

Delimitations and Limitations of the Study

This study is delimited to the Cyclical Review process piloted in nine schools during 2007–08 in one region of the Department and to the influence that the development of, the preparation for, and the participation in the reviews exerted on the eighteen participating principals over the period 2006–09.

Given these constraints the following limitations applied.

- The principals who were both team leader and host were willing and motivated participants.
- The team-member principals were willing to participate and were nominated by their school education directors as successful principals.
- The longitudinal nature of study was limited to approximately two years. The first review occurred in August 2007 and the last in November 2008, with most occurring between September 2007 and December 2007. The final data collection occurred in June 2009.

Most importantly, the study examines the influence on participating principals from three perspectives: as individuals; as leaders in their schools; and as system leaders.

While it does consider how principals perceived the influence of the review on themselves as school leaders, and therefore their perception of changed and improved practices at school level, the study does not specifically consider the influence of the review on school improvement, and especially not on that of improved student performance. Furthermore the study does not consider the influence on other review-team members.

RESEARCH DESIGN

Four recurring themes emerged from an examination of the practical context of the study and an examination of the literature regarding educational evaluation; evaluation utilisation, use, and influence; evaluation of schools; participatory evaluation; and ECB. The four themes were:

1. a model to record and analyse the data in order to determine evaluation influence
2. knowledge, prior experiences, and factors that influence principals who participate in Cyclical Reviews
3. the distinction between evaluation use and evaluation influence
4. Cyclical Reviews as ECB.

These gave rise to four research questions for this study.

1. Is there a theoretical model that can be designed to map evaluation influence in the case of Cyclical Reviews?
2. What factors, prior experiences, and understandings contribute to the influence that involvement in Cyclical Reviews in Western Sydney Region has had on the participating principals?
3. How does participation in Cyclical Reviews in Western Sydney Region influence participating principals?
4. To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?

These questions gave direction to the research design. To develop useful responses to the key research questions it was necessary to gather qualitative data on the Cyclical Review processes and influences and, in particular, the participating principals' opinions about these processes and influences. As the Cyclical Review process was designed within a participatory action-research environment (Kemmis & McTaggart, 2008:273; Punch, 2009:135–139), a qualitative approach, using a case-study technique (Yin, 2003) for the gathering of data from multiple sources, was adopted. The gathering of data thus occurred in a number of ways and at various stages. First, principals were observed by the researcher before, during, and after each review. Second, principals completed three questionnaires: the first, immediately following the review, related to the factors they perceived to have either helped or hindered the review process; the second, completed a few months after their participation, related to the immediate and medium-term influences they perceived to have occurred; and the third, completed approximately one year after their participation, related to the long-term influences they perceived to have occurred. Third, documents relating to the preparation for and implementation of Cyclical Reviews and also those from each of the participating principals' schools were analysed to look for evidence of influence. Fourth, interviews were conducted with four of the participating principals to extend the conversation begun with the questionnaires and thereby deepen understandings of principals' perceptions.

The questions also gave rise to the selection of PPODS for data collection and analysis, because it allowed data to be sorted, searched, and analysed in multiple ways. In addition, PPODS was used for data collection and analysis in the Cyclical Review process itself and therefore understood and accepted by the participants.

STRUCTURE OF THE STUDY

This study is reported in eight chapters. This chapter has outlined the purpose and scope of the study, explained the key terms to be used, and has also provided a brief description and explanation of the research design.

Chapter 2 provides a brief historical and comparative context for the study. It first provides a brief history of the evaluation of the performance of schools in New South Wales from 1842 to 2006. It identifies the antecedents of the school inspection system and provides a description of the inspection system itself that prevailed, with only relatively minor changes, for well over 100 years. It further provides a broad description of the events and arguments—identifying the historical, educational, social, and political factors—that led to the move away from school inspection to quality assurance and finally to a system of school accountability and development. Second, it provides a snapshot of contemporary school accountability and development approaches across Australia. Third, the chapter places New South Wales and then Western Sydney Region within this context, describing in detail the place of the Cyclical Review process and its implementation.

Chapter 3 is an exploration of the literature that relates to educational evaluation; evaluation utilisation, use, and influence; and the factors affecting influence. It further explores the literature relating to school evaluation and its two main functions, summative accountability and formative development, before examining in particular the concept of Cyclical Reviews as participatory evaluation and Cyclical Reviews as ECB. The research problem is presented in four themes as issues arising from the practical context of the study in Chapters 2 and the literature in Chapter 3. From these themes, the four key questions for research are identified and presented.

A short chapter, Chapter 4, develops and presents the provisional theoretical model of evaluation influence developed for the study.

Chapter 5 is concerned with methodology. The themes and research questions developed from the context and the literature, coupled with the action-research environment in which the Cyclical Reviews took place, suggested that an interpretivist–constructivist perspective was appropriate for the study. A case-study technique within a qualitative approach was used to gather data from multiple sources for analysis.

Chapters 6 and 7 report findings. They do this with an analysis of the data collected by all methods, with particular reference to recurring themes and the key questions specified in Chapter 3. Both chapters identify a range of issues arising from the analysis of data, draw implications from these issues by reference to the international literature on evaluation, and propose recommendations for Cyclical Reviews.

The final chapter is a summary of the issues, implications, and key findings of the study. It makes recommendations for future theory, system policy, and practice about Cyclical Reviews and poses questions for future research.

CHAPTER 2
THE CONTEXT OF THE STUDY

INTRODUCTION

This chapter examines the processes used to evaluate public schools in New South Wales over the period 1848–2008 and provides a snapshot of processes operating in other Australian states and territories in the period 2006–08 when the Cyclical Reviews were conducted in Western Sydney Region. This provides both a historical and comparative context, a necessary component for understanding the views of and influences on the principals who participated in this study.

The processes developed for the Cyclical Reviews in Western Sydney Region are then described in detail, thus completing, for the purposes of this study, an overview of relevant school evaluation processes for the period 1842–2008.

HISTORICAL ANTECEDENTS

The Early Period 1848–1991

Reviews have long formed a part of the educational landscape in New South Wales. From as early as 1848 school reviews, in one form or another, have been integral to the accountability function of New South Wales public school education, and elements of even the earliest reviews still remain in current review processes. Also forming an integral part of the educational landscape in New South Wales have been the voices of the New South Wales Teachers Federation and the two groups representing principals. These latter groups, the Primary Principals' Association and the Secondary Principals' Council, while not industrial bodies, have nevertheless increasingly held sway in

matters pertaining to the accountability function in schools, particularly where the supervision and accountability of principals are concerned.

A brief outline of the development of the school inspection and review system is therefore pertinent, concentrating mainly on developments in the more recent decades with which the principals in the study might have been familiar and which might therefore have affected their perceptions of and attitudes Cyclical Review process.

In 1854 William Wilkins, Superintendent and Inspector of National Schools, developed and promulgated seven areas relating to the condition of schools (Dugdale, 1990: -16) on which inspectors were to base their reports. (Wilkins's position was essentially the same as what is today called the Director-General). The seven areas were: its material state; its moral character; the subjects and methods of instruction; the proficiency of the pupils; the teachers' qualifications; the general tone of the school; and the local supervision.

> The colonial inspector was thus obligated to take direct steps to see that teachers and local managers faced up to their responsibilities and endeavoured to correct the revealed deficiencies that the inspector recorded and advised on as a result of his inspection findings.
>
> (Dugdale, 1990: -19)

While there have been many iterations of the criteria on which schools in New South Wales have been evaluated since Wilkins's time, the seven areas relating to the condition of schools can be seen as the forerunner to the terms of reference for the Cyclical Reviews.

For the remainder of the nineteenth century the inspection system developed by Wilkins remained the method by which schools and teachers were assessed. According to Dugdale (1990, 2005) there appears to have been insufficient major criticism of the system from the recipients of inspection—the teachers, both individually and

collectively—to have warranted any change in school supervision during this time; no doubt the result of the prevailing authoritarian atmosphere in which the testing of a school's and a teacher's efficiency by means of rigid inspectorial examination of pupils was generally, if not happily, accepted by those in the teaching-service.

Whereas it could be asserted that close supervision and inspection were necessary in the second half of the nineteenth century, teachers at that time having received little or no initial training, this was not the case by the mid-1950s. As aspiring teachers were required to undertake longer periods of initial training, and inservice education received greater emphasis, the perceived need for close supervision by inspectors had diminished. Thus 'the inspector's advisory function increased as school principals were encouraged to play a higher role in supervising their staff in a positive way' (Dugdale, 2005:9).

Nevertheless, the evaluation method most commonly used by the inspectorate from the end of World War 2 until the inspectorate's demise in 1989 was almost exclusively one of regular school reviews and the accompanying compulsory assessment of the performance of teachers. While principals were encouraged to supervise and mentor their staff, there was little recognition as far as official Departmental accountability was concerned of the legitimacy and value of the opinions of principals and teachers, and certainly not of the parents and students of the schools. The views of external experts remained paramount for 148 years.

The Scott Report of 1990, the subsequent Quality Assurance era, and the emergence of new school accountability frameworks, coupled with growing action by the New South Teachers Federation and advocacy by the two principals' groups, began to change this emphasis. Despite this, the views of external experts were still seen as necessary in the review of schools. It was not until the introduction of the Cyclical Review process that the recognition of principals as competent leaders of reviews was achieved.

The Scott Report recommended the abolition of the inspectorate in 1990 and in its place the establishment of a Quality Assurance Directorate to review school, region, and state programs (New South Wales Department of School Education, 1992:31).

Quality Assurance 1992–1995

Two key elements of all Quality Assurance reviews were the involvement of stakeholder groups—students, parents, and school staff—and the public reporting of findings. By mid-1995 about 60 per cent of schools had been reviewed.

Quality Assurance school reviews were conducted by teams consisting variously of principals, school executive staff and teachers from other schools, cluster directors, and parents. The leader, usually a Quality Assurance Director, was responsible for establishing the review, managing it, and producing the report, complete with findings and recommendations. The reports were public documents. The host-school principal was fully involved in the review and was a member of the reviewing team. The host-school principal and local cluster director (a senior Departmental officer with responsibility for a small group of schools) were responsible for ensuring that the recommendations of the review were implemented following the review.

There are a number of similarities between the Quality Assurance movement 1992–95 in New South Wales and the Cyclical Reviews, the main ones being: a review conducted by teams including principals, school executive and staff, and teachers from other schools; a review based on a set of guiding educational statements; and a complementary survey tool.

In 1995 the New South Wales Teachers Federation, unimpressed with proposed changes to the Quality Assurance movement, directed its members (and this included principals) to ban participation, and as a result most schools withdrew from scheduled

school reviews. The Department reconsidered its position and proposed the development of a new approach to school accountability and improvement. This approach was to take the best practices of the previous Quality Assurance initiative and add a number of features drawn from other school systems, other workplaces, and consultation with community members.

The principals involved in the Cyclical Reviews would have remembered, and in many cases participated in, the Quality Assurance movement, and they would also have remembered and possibly agreed with the New South Wales Teachers Federation's position. Certainly their representative groups, the Primary Principals' Association and the Secondary Principals' Council, were in accord with the New South Wales Teachers Federation's position. If one looks back a little further, most of the participating principals had worked in schools before the abolition of the inspectorate and their attitudes to Cyclical Reviews might also have been influenced by their memories of the inspection system. All these factors needed to be taken into account by the researcher.

It took another decade from 1995 for the concept of regular whole-school reviews in New South Wales to return to favour. In the intervening period a School Accountability and Development Framework (New South Wales Department of School Education, n.d.-a, n.d.-b), which included reviews only by exception, was introduced; this framework still remains, but with the addition of a cyclic-review component. But while this cyclic-review component exists in principle, it has not as yet been implemented statewide.

The Period 1996–2000

The paper *School Accountability and Improvement in NSW Public Schools* (New South Wales Department of School Education, n.d.-a) outlined the three-step process of evaluation and reporting to the community: self-evaluation by the school, assisted by a

Departmental school improvement officer; the production of an annual school report for distribution to parents and the wider community; and in-depth school reviews in selected schools.

Within the proposed model, in-depth reviews would be provided for schools where the information from the self-evaluation indicated capacity for a higher level of performance. These reviews would target specific areas and would focus on how student learning outcomes could be improved. In some instances high-performing schools would be reviewed to identify and share those factors that contributed to their success. The rationale provided for moving from a program of routine, scheduled reviews to one of strategically identified and focused reviews was twofold. First, it was a more efficient and effective use of resources, as it directed the Department's resources to schools most in need. Second, while schools had found the quality assurance reviews a positive process, there was little evidence of improvement in classroom practice and school processes resulting from them (New South Wales Department of School Education, n.d.-b:10–11).

At the end of 1996 the implementation of school accountability procedures was being negotiated with the New South Wales Teachers Federation and other interest parties such as the Federation of Parents and Citizens' Associations, and the principals' groups. The New South Wales Teachers Federation and the two principals' groups, wary of any accountability process, were particularly wary of the review component, especially as the premise for the majority of the reviews to be undertaken was that a school's data showed some form of underperformance.

Simultaneously with these negotiations, the Department inserted a clause (Clause 17) in the *Crown Employees (Teachers and Related Employees) Award 1996* requiring all schools to implement, amongst other things, annual reports and associated school self-evaluation and improvement programs, and school reviews. A protracted industrial dispute ensued with a ban being placed by the New South Wales Teachers Federation

on the administration of the Basic Skills Tests, school self-evaluation, and annual reporting.

It was not until 1 August 1997 that the Department and the New South Wales Teachers Federation reached agreement. This agreement 'related to the self evaluation, the structure and contents of annual school reports and the use of performance data' (New South Wales Department of School Education, 1997:62). There was, however, no agreement on the introduction of the school review component of the process.

Two years later, on 1 November, the *School Development Policy* (New South Wales Department of Education and Training, 1999) was released for immediate implementation. It addressed the remaining element of the accountability and improvement model for New South Wales government schools: school reviews. The document detailed 'policy relating to school development, viz, Education Support Teams, School Program Reviews and School Management Reviews' (New South Wales Department of Education and Training, 1999). But it was not until 2001 that the school accountability and improvement process was implemented in full and that targeted school reviews under the *School Development Policy* were conducted in large numbers across New South Wales. Even with this agreement, intense scrutiny by the New South Wales Teachers Federation was placed on the proposal for every review and the two principals' groups also voiced concern over a number of the proposed reviews.

The Period 2001–2006

The period 2001–06 is crucial to understanding the origin of the Cyclical Review process and the way it was viewed by principals. Although unintended at the time, developments in school accountability coupled with Departmental restructuring in this period provided the impetus for the Cyclical Reviews.

From the beginning of 2001 until the end of 2003 all components of the school accountability and improvement process were fully operational. In terms of the school-review component, large numbers of reviews were conducted, between seventy and eighty a year, although they never reached the anticipated numbers of 100 to 150 a year outlined in the earlier discussion paper (New South Wales Department of School Education, n.d.-a:6).

School reviews operated under the *School Development Policy*, and three types of reviews were outlined.

- Education Support Team Visits. Requested by the principal, an education support team visit was designed to recognise and develop a program or a number of programs operating in the school following consultation with those working in the program and the district school improvement officer (a senior Departmental officer with responsibility for school program accountability and improvement).
- School Program Reviews. Initiated by the district superintendent (the next iteration of senior Departmental officer with responsibilities for a group of schools) or the principal, a school program review was to occur when a program or a number of programs in the school showed the need for improvement following consultation with the principal, relevant staff, and the district school improvement officer.
- School Management Reviews. Initiated by the district superintendent, a management review was to occur when there was substantial evidence of significant dysfunction in the operation of the school, as identified by the district superintendent, following consultation with the principal and the district school improvement officer.

It was under the category of Education Support Team Visits that Cyclical Reviews were developed and piloted. This was done for two reasons. The first was that the intent of a Cyclical Review closely mirrored that of an education support team visit. The second was that by being able to develop a form of education support team visit, the process did not need further ratification by the New South Wales Teachers Federation.

As outlined by Fullan (2001) and MacBeath and McGlynn (2002), each type of review was designed to provide, to varying degrees, three dimensions of school evaluation and development.

- Support–Pressure: a continuum with high-level support at one end and strong pressure at the other.
- Bottom-up–Top-down: representing how the system sees and implements change.
- Internal–External: a continuum from self-evaluation to evaluation from an outside source.

With an education support team visit designed to develop further a program rather than address an identified need for improvement, these reviews were designed to be more supportive, to implement change from within the school, to be much more internally oriented, and to be more flexible in structure than the other two forms of review.

Nevertheless the structure and conduct of an educational support team visit tended to follow that prescribed for the other two forms of review. Once the program to be evaluated was decided and an educational support team visit was approved as the appropriate method for the evaluation, the terms of reference for the review were developed. Depending on the focus and complexity of the terms of reference and considering the availability of resources, a review team comprising approximately four to six members and led by the district school improvement officer would be determined through negotiation between the principal and the district superintendent. The team could be drawn from district office personnel; staff of the school; staff, including principals, from other schools; other district school improvement officers; district superintendents; or other personnel. The review itself would last from three to five days, during which time a range of data would be collected, analysed, and synthesised. Data collection would be conducted in a variety of ways: document analysis; student performance data analysis; interviews with staff and parents, individually or in focus groups, and with focus groups of students; classroom observation and observation of other school activities; and questionnaires or surveys of staff, parents, and students—all triangulated to confirm findings. Feedback to the

school would be provided regularly throughout the review and a close liaison would be maintained between the district superintendent, the review team leader, and the host-school principal regarding all aspects of the process. On the last day of the review a summary of the major findings and recommendations would be outlined to the school staff at an exit presentation and these would be put into a more formal report within two weeks of the conclusion of the review. In the case of an educational support team visit, the recommendations would be for internal school use.

Accompanying the policy was a detailed set of implementation guidelines and training materials for the leaders of reviews. As the reviews were scrutinised by the New South Wales Teachers Federation, strict adherence to these guidelines was enforced. If followed correctly the preparation for and conduct of the review and the development of a detailed report would be a time-consuming task. In addition to the week set aside for the review, school review leaders (usually the school improvement officer) would often set aside another week for preparation and one or two weeks to write the report.

During 2003 a major review of the Department's organisational structures occurred. A position paper, *Lifelong Learning: The Future of Public Education in New South Wales* (New South Wales Department of Education and Training, 2003) was distributed publicly and comment invited. The rationale for the proposed restructure was to create stronger and more transparent links between schools and Technical and Further Education (TAFE), to strengthen the links between planning, practice, and accountability, and to avoid duplication of roles across the system. This was to be done by reinstating regions. As a result, regions, which had given way to a district model in 1996, were re-established from the beginning of 2004.

In each region a regional director led a group of school education directors (the former district superintendents) and a team of regional support officers. In 2004 the forty district school improvement officers from the previous model, whose work had been district-based and centrally coordinated, were attached to regions to undertake one of

three functions—regional consultancy coordinator, regional data manager, and regional teaching and learning officer. The *School Development Policy* (New South Wales Department of Education and Training, 2004) was revised to reflect nomenclature changes that had occurred as a result of the Department's restructure, but there were no changes to the intent of policy itself.

With a further restructure in 2005 the number of school education directors increased to seventy-five, with each responsible for approximately twenty-five to thirty-five schools. Regional consultancy coordinators, regional data managers, and regional teaching and learning officers gained a number of these positions. The remaining regional consultancy coordinators, data managers, and teaching and learning officers were renamed school development officers and their overall number decreased to twenty-six. The school development officers' roles were determined by each regional director.

In line with this new structure and an increased accountability placed on school principals through the school education director, a new *Framework for School Development and Accountability* (New South Wales Department of Education and Training, 2005) was developed.

The framework was explained as a continuous process designed to

> ensure that schools meet the expectations of their communities. An essential part of this process is an evaluation of the extent to which a school has realised those goals. Strategies for continuous improvement are developed from the results of this evaluation.
>
> (New South Wales Department of Education and Training, 2005)

While the framework retained the three components of the earlier undated School Accountability and Development documents (school self-evaluation, annual school reports, and school reviews), it now included principals' accountability under the

school accountability component—Principal Assessment and Review Schedule (PARS), and Principal Improvement Program (PIP). Also included in the framework was a component labelled cyclic reviews, although the only note attached to it in linked web pages was the wording 'under consideration'. This became the focus, however, for the Western Sydney Region Cyclical Reviews.

The restructure had ramifications for school reviews. Many of the school improvement officers became school education directors and the role of the school development officer within regions changed depending on the various regional interpretations of the role. This resulted in a significant loss of expertise and capacity in leading reviews. There was no longer a state manager of the school review process to manage the budget and assure the quality of the process or of the finished report. The budget for reviews was divided among the ten regions according to their size. Consequently some regions had very little money for review purposes, even if they had the expertise to conduct them; others had more money but not necessarily the expertise or the desire to conduct them. As a result the number of reviews decreased noticeably. In 2004 only nineteen reviews were undertaken across the state: five educational support team visits; thirteen program reviews; and one management review.

The evolution of school-performance evaluation from an inspectorial system that reviewed schools and teachers to a broader school accountability and development framework followed a similar timeframe and pathway in the other Australian states and territories. But New South Wales differed from other states and territories in the heightened level of union opposition and unease from principals' groups regarding any form of school review. The reduced level of opposition in most other Australian states and territories possibly explains why cyclical or mandatory school reviews are a common component of their practices in the period 2005–09. The reports in the following section accordingly relate to school review provisions in the Australian states and territories over this period, the period of the development and implementation of the Cyclical Reviews in Western Sydney Region.

CONTEMPORARY AUSTRALIAN STATE AND TERRITORY PRACTICES

All government schools in Australian states and territories are subject to school accountability policies and processes, with school reviews, in one form or another, being a component of each.

> While terminology may differ, the common elements established in all Australian government school accountability processes include—a school accountability framework, school plan, school self-assessment, annual school report and school review.
>
> (Western Australia Department of Education and Training, 2006:4)

Victoria

In Victoria the *Accountability and Improvement Framework for Victorian Government Schools 2007* (Victorian Department of Education and Training, 2006:online) requires schools to evaluate progress and establish targets through an internal self-evaluation process and an external school review.

Of particular relevance to this study, because of the parallels and contrasts that can be drawn with the desk audit of the Cyclical Reviews, is the requirement that all schools undertake a self-evaluation process prior to a school review, which is then conducted either internally or externally (Victorian Department of Education and Training, 2006; Western Australia Department of Education and Training, 2006:5). Whether internal or external, the reviews focus on two key aspects. The first is student outcomes (student learning, transitions, engagement, and well-being). The second is the school performance, aspirations, and strategies in relation to these student outcomes, with an internal review also summarising progress against the school's strategic plan and its annual implementation plan (Victorian Department of Education and Training, 2006:online).

Queensland

In Queensland the *School Improvement and Accountability Framework* (Education Queensland, 2009:online) relates to planning, reporting, and cyclical reviews. A three-year planning-cycle is complemented by an annual operations plan and reporting-cycle.

Triennial school reviews are conducted, as are reviews by exception (presumably prompted by underachievement) and reviews of exemplary practice.

Some of the issues to be addressed through reviews include (Radii, 2005:33):

- learning-outcomes compared with statewide targets
- the school's response to the community's needs
- staff development and employment practices
- levels of satisfaction (students, parents, and teachers).

Western Australia

Western Australia's *School Improvement and Accountability Framework* includes a component of school reviews, which are intended to evaluate school effectiveness and provide feedback on how the school might improve its performance (Western Australia Department of Education and Training, 2008:online).

All schools undergo a Standards Review, conducted by the Director of Schools, focusing on standards of student achievement—academic and non-academic (Western Australia Department of Education and Training, 2008:12). These reviews can involve three to four visits during a review cycle, which generally occur every two years. The multi-visit, longitudinal Standards Reviews of Western Australia, conducted by a sole

official, present a stark contrast to the single-visit Cyclical Reviews of Western Sydney Region conducted by a team of principals, teachers, and a system official.

Western Australia also has four other types of reviews conducted by the Expert Review Group (Western Australia Department of Education and Training, 2008:13) that focus on:

- student achievement and improvement
- learning-experiences as responsive to student needs
- leadership and administration
- financial and resource management
- partnerships with the community.

Data drawn from teacher assessments and statewide standardised tests (in Years 3, 5, 7, and 9) have been available to school reviewers in Western Australia for a number of years (Radii, 2005:33).

South Australia

South Australia (Department of Education and Children's Services, 2007:online) requires all schools to conduct annual internal reviews (self-evaluation), and engage in a more comprehensive review every three years. The three-year review is undertaken by district directors and serves as a validation process of the site self-review.

External reviews—somewhat similar to the 2002 Quality Assurance Reviews of New South Wales—were trialled in 2002 but are no longer a mandated element of site learning plans. However, when it is believed that standards are not being met, 'Reviews by Exception' may be sought.

School reviewers in South Australia until 2006 had access to state-based literacy and numeracy standardised-test data, including mean scores over time and comparisons with like schools across the state (Radii, 2005:33). The review teams involved in the Cyclical Reviews of Western Sydney Region also had access to sophisticated student-performance data provided by the Department.

Tasmania

In 2009, the website for the Tasmanian Department of Education did not display a separate school accountability page and no reference to reviews of schools could be found. The website did, however, contain the *School Improvement Report 2008* (Tasmanian Department of Education and Training, 2009:online), which included a detailed performance report for the system and each school.

It is likely that the reports were derived from each school's four-year 'School Improvement Plan' (as described in the literature review of Australian school-accountability systems conducted by the Western Australian Department of Education and Training in 2006). At this time school reviews were conducted on a four-year cycle, with a focus on identified targets and involved principals, teachers, parents, and students as members of the reviewing teams.

Australian Capital Territory

In 2009 a new school improvement framework was released (Australian Capital Territory Department of Education and Training, 2009:online) describing a four-year cycle of school review.

> While processes, strategies and timeframes within the four-year cycle are largely managed by each school to best address their particular contexts, the

> timing of annual surveys, completion of school plans, publication of annual school board reports and external validation are generally at fixed points within the cycle. Each school will develop a strategic four-year school plan and an annual operating plan, self-assess on an annual basis and report the outcomes against this plan to the school community. Each school will also participate in external validation in the fourth year of the cycle to gain an objective evaluation of its achievements and standards of performance, and to inform future planning for continuous improvement.
>
> (Australian Capital Territory Department of Education and Training, 2009:5)

While the fourth-year validation review is described as external, the explanation of the process shows that it is to be a form of peer review with panels comprising a principal as the leader, two other school leaders, and a community member. There are clear similarities here with the Cyclical Reviews of Western Sydney Region.

Northern Territory

In 2006 Northern Territory schools were not subject to external reviews (Northern Territory Department of Employment Education and Training, 2006:online). In 2007, however, the Northern Territory Department of Education and Training released its *Accountability and Performance Improvement Framework*, (Northern Territory Department of Employment Education and Training, 2009:online) for full implementation in 2008 to improve employment, education, and training outcomes in the Northern Territory and to 'provide a system of accountability for the whole department including schools'. Planning, assessment, and reporting were to be undertaken on an annual basis and a four-yearly self-review program, based on performance profiles and other relevant data, was to be implemented. The framework also made provision for a selective audit for highly performing schools and schools in need of intervention and support.

As stated above, these reports relate to school review provisions in the Australian states and territories over the years 2005–09. It should also be noted that from 2008 all Australian schools became subject to the National Assessment Program for Literacy

and Numeracy (NAPLAN) testing, thus providing reviewers with an additional comprehensive source of data. It should be further noted that subsequent to the research timeframe of this study, some changes have been made to the review process in various states.

THE PLACE OF THE WESTERN SYDNEY REGION CYCLICAL REVIEWS WITHIN THE NEW SOUTH WALES SCHOOL DEVELOPMENT AND ACCOUNTABILITY FRAMEWORK

As discussed above, the years 2004–05 were a time of noteworthy change in the structure of New South Wales public school education. The newly established Western Sydney Region was no exception. In addition to the changes in structure, the personnel in Western Sydney changed considerably, not only at the beginning of 2004 but also between 2004 and 2005. Of the twelve senior staff in 2005, comprising the Regional Director, eight school education directors, and three school development officers, only three had been part of the senior regional team in 2004. In particular, the Regional Director had changed, and it was the newly appointed Regional Director who introduced Cyclical Reviews.

The new Regional Director, appointed at the beginning of 2005, had worked as a district superintendent in 2002–03 with a group of approximately fifty schools that were now part of the Western Sydney Region. At that time, convinced of the benefits of reviews, and especially the more improvement-oriented education support team visit that operated under the *School Development Policy*, he encouraged principals to be part of a district program of such reviews. Together with the then school improvement officer, he embarked on an ambitious program in which each school would be reviewed and the principal would take part in another school's review. Though not all reviews took place and not all review reports were completed before the 2004 restructure, a number of the principals involved commented in various regional

meetings on the value of the findings and most principals commented in particular on the value of the process itself.

With the appointment of the new Regional Director a renewed emphasis was placed on data-driven decision-making in Western Sydney schools, and with this emphasis the Western Sydney regional mantra of 'a relentless focus on learning' was introduced: measurable, improved student-learning outcomes and staff professional learning were to be the focus of all schools' improvement targets. The responsibility for school self-evaluation, student-performance data analysis, school planning, and school improvement was placed firmly on the principals. Principals were to be accountable to their school education directors, and school development officers were to provide the support required to assist principals in developing their own and their schools' capacities in these areas. The Regional Director's arrival also coincided with the increased accountability of principals under the revised *School Development and Accountability Framework*, as discussed above. The provision for cyclic reviews within this framework provided the authority for the Regional Director to develop this component within a Departmental policy.

During 2005 the work of and support provided by the school development officers, coupled with the accounts of the school education directors, revealed that many schools and school principals were either not using data-analysis techniques or evaluation instruments and methods or were not using them in a way that would result in valid and reliable data on which to base decisions, and this was particularly evident in school plans and in the evaluation sections of the annual school reports. Charged with his previous experiences as a district superintendent, his belief in the worth of school reviews, and his summation of the evaluation capacity of many of the schools and their principals, the Regional Director announced Cyclical Reviews as an initiative in the *Western Sydney Regional Plan 2006–08*.

CYCLICAL REVIEWS IN WESTERN SYDNEY REGION

The Regional Director's vision was for Cyclical Reviews to provide a mechanism that would allow each school to be judged approximately every four years against statements of exemplary practice that 'reflected school operations and performance at the highest levels' (Western Sydney Region New South Wales Department of Education and Training, 2006:1).

During the period of Quality Assurance in New South Wales (as discussed above) *Best Practice Statements* and a complementary survey tool, SchoolMap, had been developed for a similar purpose. A new draft document, *Exemplary Practice Statements* (Western Sydney Region New South Wales Department of Education and Training, 2006), was developed to build on these earlier documents and to make extensive use of ideas from policy, research, and practice. In particular, ideas contained in the Department's *Quality Teaching Framework*, *School Leadership Capability Framework*, school planning documents, *Professional Learning Policy for Schools*, and the New South Wales Institute for Teachers' *Framework of Professional Teaching Standards* were considered in the development of the statements (Western Sydney Region New South Wales Department of Education and Training, 2006:1).

The newly drafted *Exemplary Practice Statements* (Western Sydney Region New South Wales Department of Education and Training, 2006) were categorised into six domains:

- School Purpose and Performance
- Student Learning
- Teaching
- Curriculum
- Professional Practice and Responsibilities
- Management, Organisation, and Quality Systems.

Each domain contained ten statements outlining the key features of the domain. For example, the domain of student learning contained the statement: 'student learning

occurs within a stimulating, supportive, and secure environment' (Western Sydney Region New South Wales Department of Education and Training, 2006:3).

There still remained, however, the need not only for the *Exemplary Practice Statements* to be accepted by the principals themselves, but also for the concept of a Cyclical Review process to be endorsed by them. Both input and ownership by the principals were, from the outset, considered of primary importance (Wasson, 2006a, meeting 1 March).

The draft statements were distributed to the two regional principals' groups: the Western Sydney Primary Principals' Association; and the Western Sydney Region Secondary Principals' Council. A series of meetings were held with the groups from May 2006 and comments from individual principals were also invited in order to refine the statements. By the end of 2006 there was a general acceptance of the statements by the principals' groups (Wasson, 2006b, meetings with Principals' groups, 15 & 22 November).

The Cyclical Review Process

At the beginning of 2007 the Regional Director (Wasson, 2007, meeting with school development officers, 16 January) asked the school development officers to develop a proposal to commence the implementation of Cyclical Reviews as a pilot program later that year using volunteer schools and principals and based on the six domains of the *Exemplary Practice Statements*.

The verbal brief provided by the Regional Director on 16 January 2007 was to develop a process and plan that would create an evaluation partnership between the region and school principals; this would lead to greater evaluation capacity of all involved and especially of principals. The process and plan also needed to be manageable and

sustainable in terms of time, personnel, and cost. Most importantly, the process needed to be one that would conform to the agreed protocols for school reviews and the use of data, as contained in the *School Development Policy*.

To adhere to this latter requirement, it was agreed to design the process to comply with the requirements of an educational support team visit under the *School Development Policy* (New South Wales Department of Education and Training, 2004). But it was also recognised that, unmodified, an educational support team visit, with the time commitment and workload placed on the school development officer (as described above), would not necessarily deliver a manageable and sustainable process for the region's 240 schools, nor would it necessarily increase the evaluation capacity of principals.

A number of principles were therefore stipulated to guide effective implementation of the Cyclical Review process and to ensure adherence to the *School Development Policy* (2004). These were that:

- data gathering and analysis were to include document analysis (including survey), interviews, and observation
- the views of parents and students were to be sought and considered and, where relevant, parents could be invited to participate
- the six domains defined in the *Exemplary Practice Statements* were to form the terms of reference for each review
- interview questions were to be derived from the terms of reference
- team membership was to be in accordance with educational support team guidelines
- the school was to be provided with a report within an agreed and short timeframe after the conclusion of the review
- each review was to be entered into the centrally held database of statewide reviews—a database of reviews provided, proposed, and conducted under the

School Development Policy—and provided regularly to the New South Wales Teachers Federation.

To develop a sense of ownership of the process and build the evaluation capacity of many more school and regional staff and, in particular, the principals, specifications for the composition of the review team were proposed.

- A principal, rather than a school development officer, would lead each review. The principal would, however, be coached and supported by a school development officer. The team leader would be chosen by the host-school principal and in general would be chosen because he/she was a principal in the same learning community and/or was respected and trusted by the host-school principal.
- School staff, including executive staff, from within and beyond the school would be team members. Staff from within the school would be selected by the host-school principal. Staff from beyond the school would be selected from a list of approved team members. (To be approved, it was envisaged that potential team members would complete appropriate training.) The number of team members would depend on the size of the school being reviewed.
- There would be a regional consultant, selected to match the perceived development needs of the school, on each review team.
- A principal from another school in the region not closely connected to the host school would be a team member.
- Other personnel—for example, parents, personnel from special interest groups—could be members of the review team, depending on the needs of the school.
- A school development officer would be an *ex-officio* team member, acting as coach to the team leader and assisting in the review as required.

Next, to provide access to data from a range of sources for triangulation purposes, specific data-collection instruments were proposed for development and use in each Cyclical Review. It was believed that the development of specific instruments would provide a more inclusive and developmental approach and allow the collection of

sufficient valid and reliable data without overloading review teams and schools being reviewed. The instruments proposed were:

- surveys for staff, parents, and students covering the six domains from the *Exemplary Practice Statements*, administered by the school before the review
- a predetermined desk audit of documents that would provide evidence against the six domains, collected by the school in the week before the review and ready for the review team's arrival
- a small set of open-ended interview questions relating to the terms of reference that could be used by review team members in each review when interviewing staff, students, and parents
- a checklist of features to be observed during classroom visits
- a suggested list of classes and other activities to be observed on an optional basis after negotiation between the team leader, principal, and staff. (In accordance with the protocols for conducting reviews, teachers had to consent to being observed in the classroom, as reviews were not an assessment of individual teacher competence.)

In addition, the following procedures were proposed (to be accompanied by information packages, checklists, pro formas, and example documents) to assist in making the process manageable and sustainable and at the same time to build the evaluation capacity of many more school and regional staff—in particular, the principals.

- The review itself would take three days: the first two days for data collection and iterative analysis and the third day for final analysis, including the setting of future directions in collaboration with the host-school principal and school executive, the production of a report, and the presentation of the report to the whole staff.
- The host-school principal would be responsible for preparing a context statement as background for the review team and preparing the desk audit. In general, the host-school principal would coordinate the collection of the documents and observations to present his/her best practices in areas relevant to the terms of reference. The host school would also be responsible for the preparation of the

review timetable, letters to parents, consent forms for student participation, and administration and collation of the survey.

- The team-leader principal would be responsible for preparing the review team for the specific review, liaising with the host-school principal before the review to finalise arrangements, conducting the review, and finalising the report.
- In addition to coaching the team leader throughout the process, the coach would be available during the review to assist where necessary.

The Cyclical Review Implementation Plan

From their experiences in conducting over 100 school reviews, the school development officers (of whom the researcher was one) were aware that the proposed procedures and instruments alone would be insufficient to deliver a review process that would be manageable within a three-day timeframe and sustainable across 240 schools. In addition, although the school development officers had extensive training and experience in conducting reviews and were capable of developing the processes and instruments required, they were aware that this would not necessarily lead to increased evaluation capacity or ownership of the process by the principals.

A range of implementation features were proposed to make the process workable. These could be grouped under four requirements: the establishment of a working-group; the development of an electronic data collection and sorting process; the establishment of an approved and trained group of prospective team members; and the devolution of costs. These will be discussed in turn.

Working-group

The first requirement was that the principals who volunteered to take part in the Cyclical Reviews would form a working-group with the school development officers

and work throughout 2007 to finalise, trial, evaluate, and modify the process and the materials that would support it. In the preparation stage this required the group to develop a survey instrument, the interview questions, the class and school observation lists, the desk audit, and the coding keywords. Once the reviews commenced it required the group to reflect continually on the process and recommend refinements and modifications to the instruments and procedures.

Electronic data collection and sorting process

The second requirement was to develop a more streamlined approach for the collection, coding, and sorting of data, using pocket PCs, laptop computers, and spreadsheets. The accepted practice in school reviews is to have the review team member hand write information during interviews, after classroom observation, or while analysing a document onto a pro-forma page. The member then rewrites the information, after first sorting it into the categories that fit the terms of reference, onto large display sheets. These sheets are displayed round the review-team operations room and can then be seen by school staff at debriefing sessions and used by the review team to search for common themes when deliberating on findings and recommendations.

The new approach, while in keeping with the intent of accepted practice, and most importantly, ensuring that no richness of data would be lost, proposed the use of pocket PCs. Team members would use the pocket PCs to record coded findings. A new file would be created for each source of data; for example, an interview with a group of students, a document, a classroom observation, or results of the survey. Each file would be named so that the source of data could be identified, but people's identities would not be disclosed. The device would allow users to hand write directly onto the screen of the pocket PC. Each thought, sentence, or idea, called a data point, would initially be coded against one of the six domains of the *Exemplary Practice Statements* and further coded according to whether it represented an achievement or a future

direction. The domain sub-theme (keyword) of each data point would be placed at the beginning of the data-point entry. Separating the codes from the text using carriage returns enabled the data to be entered into tables and spreadsheets. Using this method the team member's handwriting could be transcribed by the pocket PC into a text file that could then be transferred electronically to a computer as a rich text file, converted to a Microsoft Word table, and entered into a Microsoft Excel spreadsheet for sorting, analysis, and report-writing. The benefits of using this technology were threefold.

1. Timesaving—data needed to be entered only once; each iteration could be managed electronically; data could be grouped and sorted in multiple ways; debriefing sessions would be always up-to-date; and data would be viewable using one or more data projectors.
2. Accuracy and reliability—data would be entered exactly as provided from each and every source and be able to be validated instantly with interviewees; no data would be lost through successive transcription; and report-writing could flow directly from the original source.
3. Triangulation—data could be grouped, sorted, and validated in multiple ways according to predetermined codes that allowed for common themes to be identified, common or diverse data sources to be identified, and outliers to be identified and discarded.

The approach was developed so that all members of the review team could collect and analyse data with only a small amount of training and practice, using software with which they were already familiar—Microsoft Excel and Word. This is the process that in time became known as PPODS. This same process was also used to collect, organise, and sort the data for the present study and is discussed in Chapter 5, Methodology.

Approved and trained group of prospective team members

The third requirement was to develop a list of prospective review-team members and to develop a suite of professional-learning activities to train team leaders and team

members in review techniques and Departmental school-review protocols, and to train school development officers (and, in the longer term, principals) as coaches. It was recognised that if Cyclical Reviews were to be fully implemented across the region there would need to be a pool of trained staff for review teams and principals for team leadership. Any staff member could express interest in being a review-team member, but staff would need the recommendation of their principal and would need to complete a training program.

As these Cyclical Reviews formed a pilot program, it was proposed to identify a small group of team members and to conduct the training on a reduced scale. The principals who volunteered would be asked to identify other staff in their schools whom they felt had the requisite educational expertise to become team members for either their own school's review or as a team member for another review. Another group of principals would be identified from across the region to take part in the reviews. The team-leader principals would be provided with intensive training by the school development officers on school review methodology, leading reviews, the use of PPODS and how to manage the technical aspects of PPODS, and the Departmental protocols for conducting reviews and using school performance data. Team members would be provided with training by the team-leader principals and school development officers on review methodology and data collection using PPODS. The school development officers would receive training in working as coaches. The role of coach would not be dissimilar to that of the 'optional senior officer,' as described in the *School Development Policy: School Review Resources* (New South Wales Department of Education and Training, 2006), 'where the team leader is relatively inexperienced in leading reviews'.

Devolution of costs

The fourth requirement was to place the cost of the review at the school level. As described above, reviews under the *School Development Policy* were expensive. A review could cost between $5000 and $8000, most of which was the cost of teacher relief; that

is, the cost of employing a casual teacher to replace the teacher who was absent from his/her normal duties. The budget for all school reviews for a year for Western Sydney Region was in the vicinity of $20 000, barely enough to cover the targeted reviews required, and not nearly enough to cover the number of Cyclical Reviews should the program become fully operational. However, with the ownership of Cyclical Reviews more firmly resting with the principals, with the focus on development, and with individual team members receiving intensive and authentic professional learning through their involvement, it was proposed that schools should cover the costs of the reviews and that the schools' professional learning funds could cover the cost of teacher relief. It was also envisaged that there would be no cost incurred by the majority of principals who took part in the reviews as they were in non-teaching roles. The region would provide the technical equipment, incorporate participation in Cyclical Reviews into the consultants' workload, incorporate time for the development and coordination of the project into the school development officers' workload, and pay for incidental expenses during the development of the project.

The resultant process and implementation plan (Ikin, 2007) were approved by the Regional Director in February 2007. Principals who had expressed interest in a Cyclical Review concept in 2006 were invited to a meeting to discuss the process. The object was to gain the principals' commitment to develop the process further, to participate in a pilot program, and to be active in the program's evaluation. From this meeting nine principals volunteered to take part.

Working-group Structure and Role

The working-group comprised the nine volunteer principals and the school development officers. The principals had worked in the district that underwent a review process in 2002–03 and were positively predisposed to the concept of Cyclical Reviews. They also held views about what could be improved, the most cited being the need for the school to receive a report as soon as possible after the review.

The working-group met in April 2007 and again in May 2007. At the first meeting the group was given a practical demonstration of PPODS and made final adjustments to the wording of the *Exemplary Practice Statements*. The purpose of the second meeting was to discuss the proposal in detail and assign tasks to individual members. At the meeting on 9 May 2007 subgroups of three to four principals were formed to develop:

- a Likert-type-scale survey instrument
- a required list of documents for the desk audit
- a set of interview questions
- a suggested observation list
- codes and keywords for recording and analysing data using PPODS.

To ensure consistency between the instruments, principals formed part of at least three subgroups. In addition, one of the school development officers (the researcher) chaired the working-group and liaised with each team to provide technical assistance and advice in evaluation methodology.

The school development officers, in partnership with one of the working-group principals, developed the two training programs—for principals and team members—in the use of PPODS and review methods and the support packages for the host-school principals and team-leader principals, including pro-forma letters, consent forms, an example timetable, a report template, and checklists for team leaders and host-school principals. With the exception of PPODS, this work was based on the materials contained in the *School Development Policy: School Reviews Resources* (New South Wales Department of Education and Training, 2006) and the school development officers' experience in and knowledge of school reviews. In line with the principals' concerns, a review report template was also developed to simplify and speed up reporting. This had not been considered in the initial proposal.

In May 2007 the school development officers completed a coaches' training-course. From July to September 2007, the subgroups developed the instruments using emails and face-to-face meetings. During this time the principals negotiated (each with

another principal) to lead their reviews, nominated dates for their reviews, and recommended various members of their staff to take part in their own or another review. Commencing in July 2007, the school development officers conducted training with the principals in PPODS and in leading school reviews. They also sought nominations from the regional school education directors whose schools were not being reviewed to identify principals whom they believed would benefit from or contribute well to a review as a team-member principal. Training for team members, including team-member principals, commenced in August 2007. Specific training for each review led by the team-leader principal occurred within two weeks of each review.

The working-group met in late August 2007 to review and make final adjustments to the draft instruments and timetable for the reviews and again in November 2007 to assess the reviews that had already taken place, to examine the evaluation data already collected, and to suggest refinements and modifications to the instruments and process.

The Reviews

The reviews took place between early September 2007 and mid June 2008, with two high schools and seven primary schools being reviewed. As previously discussed, three principals participated in each Cyclical Review. The first, the principal of the school being reviewed: the host-school principal. The second, the principal leading the review: the team-leader principal. The third, the principal from another school: the team-member principal. The selection of these principals, the frequency and incidence of the reviews, and issues relating to methodology are discussed in Chapter 5.

Steering Committee Structure and Role

In 2008 there was positive feedback from the reviews. The Regional Director believed that 'the pilot [was] proving highly successful' (Wasson, 2008, pers. comm. 12 February) and he established a Cyclical Review Steering Committee 'to oversight and drive the future direction of the process' (New South Wales Department of Education and Training Western Sydney Region, 2008).

Specifically, the committee, led by the Regional Director, was to

> assist in consolidating the program and procedures, ensuring validity, consistency and integrity, overseeing the ongoing development of robust instruments and review outcomes, and monitoring the program's implementation and quality. Working groups of principals with the assistance of a school development officer will continue to develop and modify instruments as required following decisions of the Steering Committee.
>
> (Wasson, 2008, letter of invitation to Steering Committee, 12 February)

The committee composition reflected the need to continue to build principals' commitment to and ownership of the process, while at the same time ensuring wider representation and expertise from across the region. It comprised:

- the Regional Director, as Chair
- a school development officer as executive officer (this researcher)
- two primary principal representatives—one from the regional Primary Principals' Association and one from the pilot schools
- two secondary principal representatives—one from the regional Secondary Principals' Council and one from the pilot schools
- a representative from the regional Council of Parents and Citizens' Associations
- a university representative to assist in research and instrument design
- two school education directors
- the other two school development officers

- the Leader, Regional Professional Learning, to assist in the development of training packages and the development of a website, and to coordinate regional consultants' participation.

At the Steering Committee's first meeting on 2 April 2008, the Regional Director commented that

> the pilot has been extremely successful and a major factor in the success to date has been the ownership that has been accorded to the principals from the participating pilot schools. Review team members and principals and staff from reviewed schools have provided positive feedback about the process (including the instruments developed) and the outcomes of the reviews. They have also provided constructive feedback on modifications that will make the process more robust. Importantly the work commenced in examining the use principals are making of the reviews—the use of the reports and the broader influences the reviews are having on principals' professional learning as well as that of their staff—is very encouraging.

Nevertheless, some reservations were voiced, especially from the principals who had not participated in the reviews. For example:

> [I have] no difficulty with the concept as outlined as long as it is not mandatory.

> [I] would be concerned if this was yet another framework principals were being asked to operate under. In this regard the *Exemplary Practice Statements* need to have coherence.

> A number of principals who may have reservations are basing this on past experience with other forms of reviews. We need to provide them with information on this process.

> Sustainability may also be an issue—lots of programs have come and gone.

> There should be a connection between the reviews and regional support but we need to be careful of the balance between a development and accountability function.

Despite these reservations, wider implementation of Cyclical Reviews for 2008 gained the full support of the committee and, most importantly, the two regional principals' groups. During 2008 the Steering Committee endorsed the ongoing collection of data from principals to evaluate the review process and acknowledged the influence the reviews were having on participating principals. From these data the Steering Committee requested working-groups to make modifications to the review instruments and oversaw reviews in eight more schools using the revised and modified instruments and processes. The Steering Committee endorsed Cyclical Reviews for full implementation in Western Sydney Region from mid-2009.

SUMMARY

This chapter began by examining the various review processes that have been used in New South Wales government schools since 1848, showing a long history of evaluation by external inspection. The chapter then showed that since 1991 the focus has changed fundamentally from an external (to the school) inspection model to a more collaborative system of review, involving principals, teachers, and community members, as well as Departmental officers. In providing this overview, the chapter also highlighted some of the prevailing opposition to various forms of school accountability and evaluation that has resulted in a deep-seated wariness by some that even today surrounds school accountability, evaluation, and reviews in New South Wales government schools.

Similarly to the New South Wales government system, various types of school reviews have formed and continue to form an integral component of all Australian states' and territories' school accountability and development frameworks and each of these frameworks has been described.

Finally, this chapter has explained in detail how the Western Sydney Region Cyclical Review process was developed during 2006–07 and piloted in nine schools during 2007–08. Cyclical Reviews, while grounded in the *School Development Policy*, were shown to rely on specifically developed evaluation instruments, review procedures, data-collection and data-reporting devices, and designated team membership. In doing so, this chapter has described how Cyclical Reviews were developed to encourage ownership of and commitment by the principals to the review findings and process through a participatory approach and at the same time to build the principals' skills in evaluation.

Chapter 3 examines the literature and research that is related to the above context. It develops a definition for this study of the term *educational evaluation*, defines terms related to evaluation use, arguing as appropriate for this study the use of the term *evaluation influence*, and positions Cyclical Reviews within participatory and capacity-building evaluation frameworks.

CHAPTER 3
REVIEW OF LITERATURE

INTRODUCTION

This chapter first provides a review of the literature that is related to educational evaluation, evaluation utilisation, use, and influence, evaluations of schools, participatory evaluation, and evaluation capacity building (ECB). After developing an appropriate definition of educational evaluation for this study, the chapter analyses the definitions of terms related to evaluation use and presents them in approximately chronological order of their development. This analysis will show that these terms are limited in their applicability to this study.

Second, the chapter provides and discusses two major and more recent frameworks used to define and map the broader concept of evaluation influence, and discusses the factors that may trigger evaluation influence.

Third, the chapter examines the functions and forms of school evaluation before positioning Cyclical Reviews within participatory evaluation and capacity-building evaluation frameworks. Within these sections, the four key research questions that emerge from this review are posed as they occur.

Finally, and, with reference to pertinent contextual factors, the four research questions are synthesised under four themes. The first is a model to record and analyse the data. The second is knowledge, prior experiences, and factors that influence principals. The third is the distinction between evaluation use and evaluation influence. The fourth and final theme is about Cyclical Reviews as ECB.

DEFINITIONS OF EDUCATIONAL EVALUATION

Defining educational evaluation has drawn on a rich history of scholarship in this area. The purpose has been to refine a definition so that it is appropriate to the context of the research (ontology), is compatible with the form of knowledge required to address the problem (epistemology), and justifies the methodology required to answer the research questions (Cooksey & McDonald, 2011:186–187).

Tyler (1950:69) provided an early definition of educational evaluation as the 'process of determining to what extent educational objectives are actually being realised'. Alkin (2004a:3) recognised Ralph Tyler's works as providing many of the bases for contemporary approaches to evaluation. For example, Stufflebeam (2004) defined evaluation as providing information for decision-makers, as a process that draws conclusions about an evaluand's merit or worth, and as 'an essential concomitant of improvement and accountability' (2004:262). Cronbach (according to Greene, 2004:178) sought an engaged, influential role for evaluation, and so he framed evaluation as a fundamentally educative endeavour. The Joint Committee on Standards for Educational Evaluation (1994:3) defined evaluation as 'the systematic investigation of the worth or merit of an object'. Nevo (2006:442) defined educational evaluation as the process of 'collecting systematic information regarding the nature and quality of educational objects', a definition that combines both description and judgement but distinguishes between them. Rossman and Rallis (2000) regarded evaluation as learning, which is itself a socially constructed appreciative process.

These concepts lead to evaluation's being defined within the context of what evaluators do; namely, undertaking evaluative inquiry: responding to a range of decision-makers' information of which establishing the worth of a program is just one (Owen, 2006:17). Thus Owen (2006:18) provided an encompassing definition of evaluation as 'the production of knowledge based on systematic enquiry to assist decision-making about a program'. Correspondingly, the Cyclical Review process is

based on the *School Development Policy* (New South Wales Department of Education and Training, 2004:1) which states that

> an effective school will be one that is constantly striving to enhance its educational provision through a process of self-evaluation, reporting and development within the resources available to it [and where in particular objective 1.1 is] to assist the school to develop a program within the school by a team that works with the school to recognise and develop a program or a number of programs operating within the school.

Owen (2006:18) included as a subset of his definition the need for 'evaluation as the judgement of worth of a program'. The Cyclical Review process (New South Wales Department of Education and Training Western Sydney Region, 2008) strengthened the School Development Policy by incorporating a judgement-of-worth component through a requirement to evaluate school programs against a predefined set of *Exemplary Practice Statements*. The match between Owen's definition of evaluation and the intent of the Cyclical Review process is therefore evident and was used throughout the research reported in this study.

DEFINITIONS OF EVALUATION USE

In his chronology of evaluation practice Alkin (2004b) suggested there were three distinguishable branches of evaluation theory: evaluation methods; valuing; and use. As this third branch, use, is the focus of this study, the development of a definition of the term *use* is necessary and is drawn from the work of those theorists fitting Alkin's classification of use. To the extent that they contributed to evaluation influence, the work of theorists who Alkin placed on the evaluation methods branch (for example, Cronbach, Weiss) and valuing branch (for example, Guba, Scriven) is discussed later in this chapter, but further discussion of their work is irrelevant in this section where the focus is on the definition of use.

In Alkin's (2004b) classification of evaluation theorists' views and influences, he firmly planted theorists such as Stufflebeam, Wholey, Patton, Alkin, Fetterman, Cousins, King, Preskill, and Owen on the use branch. Nevertheless, although some in the evaluation profession challenge the primacy of use (Henry, 2000; Henry & Mark, 2003; Scriven, 1991), many, if not most, evaluators agreed that the use of evaluation was an issue of prime importance (Preskill & Caracelli, 1997).

There is much literature related to the terms *evaluation utilisation* and *evaluation use*. Indeed evaluation utilisation and use have been major concerns of theorists, researchers, evaluators, and stakeholders since the 1970s. Shulha and Cousins (1997:196) credited work in the 1970s (for example, the work of Alkin, Daillak, & White, 1979; Alkin, Koescoff, Fitz-Gibbon, & Seligman, 1974; Patton et al., 1977; Weiss, 1973, 1979, 1977) as 'significantly shaping evaluation utilization as a field of inquiry'. Christie (2007:8) noted that 'evaluation utilization is arguably the most researched area of evaluation'.

The generic dictionary meanings of the terms *utilisation* and *use* are almost identical. *Utilisation*, as defined, for example, in the *Australian Concise Oxford Dictionary* (1997), means 'to make practical use of, turn to account, use effectively'. That dictionary defines *use* as 'the act of using or state of being used, application to a purpose, the ability to be used, [and] the purpose for which a thing can be used'. Fowler (1965:670) notes that 'if differentiation were possible between *utilize* and *use* it would be that *utilize* has the special meaning of make good use of, especially of something that was not intended for the purpose but will serve. But this distinction has disappeared beyond recall'.

Applied to evaluation, the meanings of the terms *utilisation* and *use*, while essentially retaining this generic meaning, have nevertheless been delimited. At the same time, the meanings have also evolved considerably. In the 1970s the term *utilisation* was used

almost exclusively and was defined simplistically as relating to the direct use of evaluation findings (Johnson, R. 1998).

In the 1980s, *utilisation* was also referred to as *use* and with the language shift came a more complex definition. By the end of the 1980s, the evaluation field was using a 'three-faceted concept of evaluation use' (Lawrenz, Gullickson, & Toal, 2007:276), with *instrumental use*, *conceptual use*, and *symbolic use* as its categories (King & Pechman, 1984; Leviton & Hughes, 1981), 'although there were minor variations in categorization and different authors used somewhat different terminology' (Weiss, Murphy-Graham, & Birkeland, 2005:13). Furthermore, Owen (2006), Cousins and Leithwood (1993), Greene (1988), and Huberman (1987) showed that these three types of evaluation use were all interrelated, that all three types could exist in a single evaluation, and that conceptual use tends to precede instrumental and symbolic use. As R. Cummings (1999:2) so eloquently described

> That is, an individual usually needs to have information which assists in better understanding a program or the issues related to it prior to being able to use study findings for either instrumental purposes (making a decision) or strategic purposes (mounting an argument or influencing others).

What these three types had in common was that they resulted from the findings of an evaluation. They have accordingly been termed as either *findings use* (for example, Alkin & Taut, 2003; Hofstetter & Alkin, 2003; Patton, 1997) or *results-based use* (for example, Kirkhart, 2000). In the 1990s a broader concept of use, adding *process use* to findings use/results-based use emerged (Alkin & Taut, 2003; Cummings, R., 2002; Johnson, K. et al., 2009; Lawrenz et al., 2007). Over time, other categories of use have been proposed; for example, *legitimative use* (Owen, 2006) and *imposed use* (Weiss et al., 2005). Mark and Henry (2004:39) believed that the term *use* had become overgrown in some senses and impoverished in others: overgrown in the sense that there are now a number of overlapping terms and impoverished in the sense that the change processes that lead to social betterment are missing from the literature. The concept of *misuse* (for

example, Alkin et al., 1979; Muscatello, 1988; Owen, 2006; Weiss et al., 2005) is also discussed in the literature.

The following sections examine the abovementioned terms (*instrumental use*, *conceptual use*, *symbolic use*, *process use*, *imposed used*, and *misuse*), all of which have been used in the literature to define various aspects of the concept of use. The discussion is important for determining the adequacy and appropriateness of these terms for the requirements of this study.

Instrumental Use of Evaluation

Dominating early evaluation-utilisation research (Alkin et al., 1979), *instrumental use* refers to the immediate direct use of evaluation findings to affect a decision or action about the program under evaluation (Weiss, 1977; Weiss et al., 2005). The interest in instrumental use arose from the need to justify evaluations, maximise the positive social impact, and address concerns over perceived non-use of results (Ciarlo, 1981). Thus instrumental use is defined in terms of the type of change, with regard to the actions that happen as a consequence of evaluation (Mark & Henry, 2004:36).

Patton referred to this use as 'intended use by intended users' (1978, 1997). In this definition *intended use* can only be use of the results and intended users can only be those directly affected by the results: the decision-makers. Any other type of use would be unintended. The problem with this is that a limitation of *instrumental use* to intended use by intended users precludes the possibility of unintended use by (intended or unintended) users. Of the four combinations, intended and unintended use by intended and unintended users, Patton's definition covered only one, namely intended use by intended users. In this study, to ignore three of the four combinations could miss potentially valid outcomes. Kirkhart (2000:12) expressed similar reservations when she noted that historically *instrumental use* meant that intended users were

identified early in the process with the express purpose of shaping each facet of the evaluation process to maximise the use of the results.

Kirkhart's view that the historical meaning of the term *evaluation use* was limited to intended use by intended users was borne out by the interpretation of the term in many empirical studies (for example, Patton et al., 1977; Rich, 1977; Weiss, 1979, 1980) undertaken during the 1970s and early 80s to discover how users or prospective users made use of the evaluation findings. Larsen and Werner (1981), in their study measuring use in a mental health program consultation, defined four types of use, all of which related to the direct use of the results. They were 'complete implementation', 'adaptation of information', 'partial use of information', and some measures 'taken toward implementation'. In these early studies it was assumed that evaluation findings had a direct and linear relationship with decision-making processes (Patton, 1978; Weiss, 1977). As a result the early research (especially research in 1970s), by virtue of its emphasis on summative use of results and its definition of *utilisation*, which was 'too narrow and failed to take into consideration the nature of actual decision-making processes in most programs' (Patton, 1997:7–10), established a pattern of finding weak links from evaluation results to program decision-making, underuse, or non-use (Alkin et al., 1979; Patton et al., 1977; Weiss, 1982a).

Conceptual Use of Evaluation

At the end of the 1970s and in the early 1980s the concept of *use* expanded. Notably, Weiss (1980, 1981, 1982a) was an early proponent of *use* being expanded to include *enlightenment* (as termed by Weiss, 1980) or *conceptual use* (as termed by Greene, 1980; Patton et al., 1977). Using decision theory and organisational psychology, Weiss (1982b) argued that *use* should be redefined to include a broad range of activities, including more subtle and gradual processes such as greater understanding and changes in attitudes by policy-makers about a particular policy or program. In doing so, Weiss (1981) further challenged the appropriateness of the term *utilisation*, arguing

that it had restricted connotations of instrumentality and episodic application. In its place she suggested the term *use*.

As it became more widely acknowledged that the direct use of evaluation results was not the only way in which an evaluation could be seen to be used, a definition of conceptual use took shape. Patton (1977), Alkin et al. (1979), and Weiss (1981) had all acknowledged the complex and changing nature of policy decision-making, and they and others (for example, King & Thompson, 1983; Leviton & Hughes, 1981) had also examined the incremental and cumulative nature of such decision-making. The function of conceptual use was seen as 'educative' (Shulha & Cousins, 1997:196) and the definition therefore evolved to mean that use had occurred if, as a result of evaluation findings, decision-makers changed their thinking or attitudes, developed awareness, or increased their knowledge or understanding of the program, policy, or the issues related to the policy or program (Cousins & Leithwood, 1986; Cummings, R., 2002; Lawrenz et al., 2007; Mark & Henry, 2004; Weiss, 1981). With conceptual use, however, although people's understandings are affected, no direct or immediate action is necessarily taken (Johnson, K. et al., 2009:378; Lawrenz et al., 2007:276).

Conceptual use is therefore worthwhile, but the view taken by the researcher is that even though the decision-maker might not act upon the findings immediately or directly, he or she or others may do so later or indirectly: any of these subsequent actions should also be considered an evaluation use.

Symbolic Use of Evaluation

In addition to instrumental and conceptual use, the term *symbolic use*, also referred to as *persuasive use*, *political use*, and *strategic use*, emerged as a third type of results-based use and was seen to have a political function (Cummings, R. 2002; Leviton & Hughes, 1981; Owen, 2006; Shadish, Cook, & Leviton, 1991).

There appear to be a number of nuances to the definition of this type of use. In general, the term refers to the use that can be made of the evaluation findings in terms of justification. Thus symbolic use occurs when people take advantage of evaluation findings for perceived personal gain, to persuade others, or to use evaluation findings to gain particular outcomes, such as convincing others to support particular positions or defending a position from attack (Alkin & Taut, 2003; Kirkhart, 2000; Shulha & Cousins, 1997; Weiss, 1998a). Alkin and Taut (2003) further suggested that evaluation findings may be used to raise an organisation's reputation as one that is willing to address accountability. K. Johnson et al. (2009) suggest that when the 'mere existence of the evaluation, rather than any aspect of its results' is used to persuade or convince, symbolic use occurs.

Symbolic versus legitimative use of evaluation

Owen (2006) differentiated between *symbolic use* and *legitimative* use. He saw symbolic use quite narrowly as 'consistent with an evaluation being commissioned by an individual or a group that has no interest in applying the results [to the extent that] no modifications to the existing program follow from the dissemination of evaluation findings' (p. 111). He included such examples as 'going through the motions to satisfy the need for an evaluation as part of a contract', or commissioning the evaluation to 'curry favour with a superior', for 'career advancement', or to 'include on a *curriculum vita* [sic]' (p. 111). When findings from an evaluation are used for legitimate purposes to justify decisions already made about a program, Owen (2006:110)—in recent years—has separated this from *symbolic* use, referring to it as *legitimative use*.

The concept of differentiation of the various types of symbolic use is worthwhile for this study because of the principals' motives for participation and subsequent actions. For example, the principals of the host schools all expressed a desire to have their schools evaluated and to act on the findings. How genuine this desire was needed to be determined by the researcher. But the motives for participation and possibly

subsequent actions could be more complex than the above classification can handle. For example, the genuineness of the desire for evaluation of the host-school principals could require evidence accumulated over a long period, whereas results-based use refers to immediate action by the decision-maker. As a further example, if a principal on a review team applied the findings to justify a program in his or her own school, this might seem to be loosely a case of either symbolic or legitimative use (precisely which one would need further determination), but to classify it formally under these terms would require a broadening of the definitions. Furthermore, while the terms *symbolic* or *legitimative* are purely results-based, it is this researcher's contention that other process-based decisions could demonstrate symbolic or legitimative use.

Imposed use of evaluation

Weiss, Murphy-Graham, and Birkeland (2005) proposed another main way in which evaluators exert an influence, adopting the term *imposed use*. They saw imposed use occurring when an organisation is obliged to pay attention to evaluation findings because of a mandate or incentive given by a superordinate body.

The study by Weiss, Murphy-Graham, and Birkeland investigated how evaluations affected the Drug Abuse Resistance Education (DARE) program. At the time many school districts in the United States of America were implementing DARE, although a number of evaluations did not show that DARE sustained changed attitudes or reduced students' drug-use. The Safe and Drug Free Schools office of the United States Department of Education then introduced a requirement that to qualify for federal funding, school districts needed to show scientific (or evaluative) evidence that any program they implemented was reducing students' use of drugs. As Weiss, Murphy-Graham, and Birkeland found, districts were not prepared to undertake evaluation studies themselves, and although there was no delimiting prescribed list of approved programs, many districts construed the requirement to mean that they had to choose a

program appearing on the published list of approved programs in order to receive funding. As a result, DARE was discontinued or scaled back in many school districts.

Weiss, Murphy-Graham, and Birkeland (2005:25–26) believed this type of imposed use might become an increasingly widespread and common form of use, for two reasons. First, many current evaluations set out to study the most successful site for a particular program. 'Such a purpose for evaluation almost invites pressures to adopt the best practices identified' (p. 26). Second, the now widespread emphasis on accountability called for 'systematic evidence as the basis for judgements' and provided evaluation with 'enhanced clout' (p. 26).

Weiss, Murphy-Graham, and Birkeland (2005:26) suggested that this was just another form of instrumental use, because it used the evaluation findings to make a decision. Their study also found, however, that in many cases the districts were not at all interested in the evidence. Instead they were responding to the federal mandate. They argued that this was not instrumental use but a 'straightforward imposition'.

The researcher argues that this is not another type of use but rather, depending on the context, an example of either instrumental use or symbolic use. If the decision-maker or organisation can genuinely be seen to use the evidence to make an informed decision about which program to adopt, the researcher would categorise this as instrumental use; that is, direct use of evaluation findings (Alkin et al., 1979). If, however, the decision-maker or organisation makes no use of the evaluation findings but is merely persuaded to adopt a program owing to a mandate, use has fulfilled a political function. Therefore this type of use fits into the category of symbolic use. The factor or precursor to use—in this case a mandate—is, however, relevant and is an important factor to consider in the context of the current study, because the motives for the principals' actions are considered by the researcher to be as important as the actions themselves in understanding the influences of an evaluation on participating principals.

Evaluation Misuse and Pseudo-evaluation

Misuse and *pseudo-evaluations* are further terms associated in the literature with results-based use. According to Alkin, Daillak, and White (1979), intention to use marked the boundary between use and misuse. Weiss et al. (2005:13–14), Muscatello (1988:30) and Owen (2006:110) suggested that when decision-makers use evaluation evidence to support or justify what is in place or what they want to do, this is a legitimate symbolic use of evaluation findings. According to Weiss et al. (2005:14), however, if decision-makers bias, distort, or omit significant elements of the findings to justify what is in place or to do what they want, then misuse occurs. Others, for example Leviton and Hughes (1981) and Stufflebeam (2001), argued that the term *evaluation* should exclude activities that use findings selectively to manipulate others, delay decisions, or project a distorted image, or at least the term *pseudo-evaluation* should be applied. Cousins and Shulha (2006:281–282) provided yet another framework for classifying 'intended user uses and misuses of evaluation findings'. Theirs was a four-quadrant framework that divided use into legitimate use and misuse, and non-use into justified non-use and unjustified non-use.

The proliferation of terms to describe different aspects of instrumental use—*symbolic use*, *legitimative use*, *legitimate symbolic use*, *imposed use*, *non-use*, *misuse*, *pseudo-evaluation*, *justified non-use*, and *unjustified non-use*—while meaningful when restricted to one author, becomes overcomplicated and confusing for the purpose of this study. What this study will determine is whether the findings of a Cyclical Review were used or not used by the participating principals and whether that instrumental use or non-use was justified/legitimate or not justified, and what *influence* this *use* had. These considerations will be incorporated in the research framework.

Process Use of Evaluation

Results-based use—that is, instrumental use, conceptual use, and symbolic use—addresses the use of evaluation findings. While instrumental, conceptual, and symbolic use all have some applicability to this study, each has been shown to have limitations. And even if these limitations are addressed, results-based use alone is insufficient to determine the influence of the principals' participation in Cyclical Review process.

Consequently, another type of use is required. In its earliest appearances in the evaluation literature a fourth use, process use, emerged as a way of facilitating results-based use (Kirkhart, 2000:9). It was not until the 1990s, however, that *process use* became part of the *use* lexicon, following Patton's (1997) work on the subject. He (1997:88) defined *process use* as the 'ways in which being engaged in the processes of evaluation can be useful quite apart from the findings that may emerge from those processes'. He identified four primary types of process use (pp. 90–91):

1. enhancing shared understandings
2. supporting and reinforcing the program
3. increasing engagement, self-determination, and ownership of participants
4. program and organisational development,

thus suggesting that process use occurred when individuals change their thinking or behaviour or when organisations change their procedures and culture as a result of the learning that occurs during the evaluation process.

In 1988 Greene proposed a three-dimensional model of 'process-based influence', suggesting that participants' experience in an evaluation study can be cognitive, affective, and political. Greene's cognitive dimension referred to changes in understandings stimulated by discussion, reflection, and problem analysis as an integral component of the evaluation process. She linked the cognitive dimension to conceptual use; for example, when reflection leads to a decision or action. The affective dimension referred to the individual and collective feelings of self-worth and value

that result from being involved in the evaluation process. The political dimension referred to the evaluation process and its ability to create new dialogues, draw attention to social problems, or to influence power relationships.

Learning-to-learn at the individual, group, or organisational level, developing networks, boosting morale (Forss, Rebien, & Carlsson, 2002), developing organisational culture, and developing organisational learning (Cousins, 2003; Torres & Preskill, 2001) have also been proposed as examples of process use.

In relation to process use, Preskill, Zuckerman, and Matthews (2003:424) proposed that

> if stakeholders do indeed learn about evaluation, about the program being evaluated, and about each other from their engagement in the evaluation process, it can be said that individual learning has taken place. However, if they share their learning with others in the organisation, it is conceivable that team and/or organizational learning may occur as well.

Further, they viewed process use as reflecting social constructivist learning theory because

> it focuses on how groups of people make meaning as they conduct an evaluation. By encouraging dialogue and reflection, and by questioning assumptions, values, and beliefs, individuals come to more fully understand the evaluand, the organization, themselves, each other, and evaluation practice.
>
> (Preskill et al., 2003)

Defining process use as a fourth and parallel category of use has caused some debate. While this definition appears to be accepted by many, some have suggested that process use is not a new kind of use but rather a new source of use. Weiss et al. (2005) took this view, as did Alkin and Taut (2003), who suggested that process use might promote instrumental, conceptual, and symbolic use. In the past ten years Kirkhart (2000), Henry (2000, 2003), Henry and Mark (2003), and Mark and Henry (2004) have also defined process use as a different source of use but at the same time have

reconceptualised it. The applicability of their definitions and frameworks to this study will be discussed below.

Limitations of the Term *Evaluation Use*

As can be seen from the above discussion, the evolution of the concept of evaluation use and the resultant definitions of the terms *utilisation* and *use* have been refined by an increasing recognition of the complexity and multiple attributes of use. Nevertheless throughout this evolution, and despite suggestions (discussed above) that various dimensions of use may be interrelated, the traditional definitions still suggest a collection of separate effects that tend to occur at the time of or immediately after the evaluation. Thus some evaluators (for example, Alkin, 2003; Hofstetter & Alkin, 2003) have restricted *evaluation use* to the specific environment and general timeframe in which the evaluation was conducted, while others have treated the domain of evaluation use more broadly, including changes that occur outside the original environment' (Mark & Henry, 2004:39). This study looks at evaluation more broadly than is possible with either of the above two domains, although it is closer to the interpretation described by Mark and Henry.

This discussion leads to the tentative conclusion that the traditional theories and definitions of *evaluation use* displayed limitations that impeded their ability to interpret fully the impact of Cyclical Reviews on participating principals. Theories and definitions of *evaluation influence* are now emerging. These are discussed in the next section, where a justification of their relevance to this study is provided.

DEFINITIONS OF EVALUATION INFLUENCE

A number of researchers (for example, Henry & Mark, 2003; Kirkhart, 2000; Mark & Henry, 2004) have come to view evaluations as having both direct and indirect influences on individuals, programs, and communities. Focusing solely on the direct use of either evaluation results or processes has, in their view, not captured these broader-level influences adequately.

The recognition of the need to consider the broader-level influences has resulted in a call by the abovementioned researchers for the term *influence* to be adopted into the terminology. An interesting issue is whether the term *influence* is necessary or sufficient or both, generically and conceptually, to describe adequately the effects of the Cyclical Review process on the participating principals and whether this term is more appropriate than the terms *use* and *utilisation*. This issue is addressed below.

For a generic definition of *influence*, this study will adopt that of the *Australian Concise Oxford Dictionary (*1997*)*; that is, 'the effect a person or thing has on another'. Thus the dictionary definition of *influence*, compared with the definitions of *utilisation* and *use*, differs markedly, although it could be argued that *influence* includes the meanings contained in the other two terms and then goes further, neither limiting nor delimiting what that effect may be. In this sense, i*nfluence* is a much more subtle and broad term. From the generic point of view the use of the term *influence*, as here defined, not only captures the conceptual definitions of *use* and *utilisation* but also caters for the broader and more subtle affects that this study attempts to interpret. The remainder of this section examines the conceptual definition of influence and its necessity and sufficiency for this study.

Kirkhart's Definition of Influence and Integrated Theory of Influence

Kirkhart (2000:5–8) argued that with the growth of the evaluation profession earlier discussions and understandings about the impact of evaluation became disconnected and fragmented. She argued that the concepts of utilisation and use and the terms themselves are no longer adequate. Her concerns are based on the following.

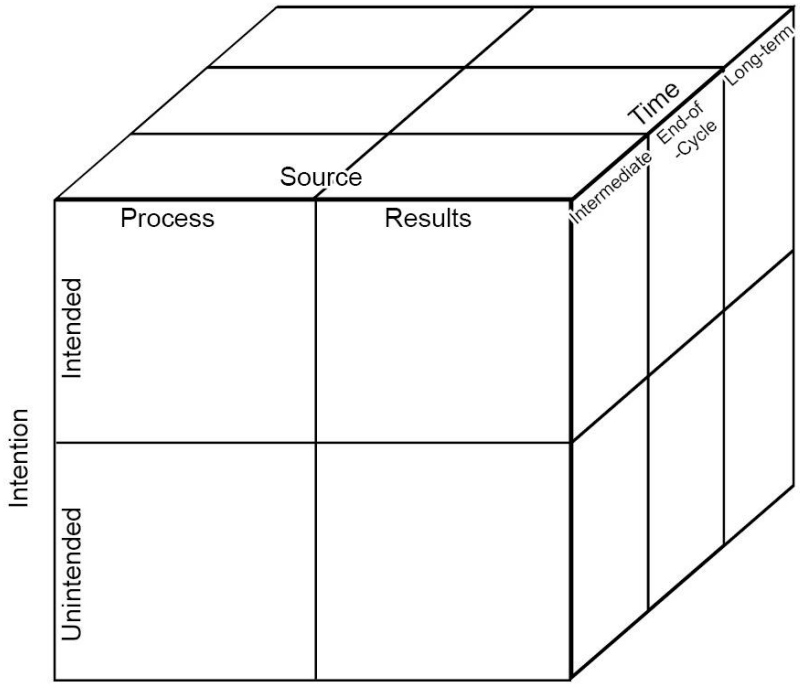
- Thinking on use lacks a theoretical perspective and limits the concept to 'unidirectional, episodic, unintentional, and instrumental' use (p. 7).
- The term *use* does not adequately describe non-results-based impact, unintended results, or incremental impact.
- The term *use* is too narrow, because it remains associated with its historical context of data-driven social research.
- Uses other than results-based are seen as secondary, or important only to the extent that they assist results-based use.

In raising her concerns, Kirkhart (2000:7) suggested that the question that should be asked about evaluation impact was, 'How and to what extent does evaluation shape, affect, support, and change persons and systems?'. To do this she argued a purposeful shift in terminology—from *use* to *influence*—and defined *influence* as 'the capacity or power of persons or things to produce effects on others by intangible or indirect means' (Kirkhart, 2000:7). This definition has narrowed the definition of *influence* by excluding tangible or direct uses of evaluation as a means of influence.

In order to examine these effects Kirkhart proposed an integrated theory of influence based on the construct of influence interacting across a three-dimensional matrix with dimensions of source, intention, and timeframe, as shown in Figure 3.1. She asserted that the model had the capacity to examine multidirectional, incremental, unintentional, and non-instrumental effects alongside the traditional unidirectional, episodic, intended, and instrumental effects. This seems to contradict her definition of *influence*, as these latter effects are those that have traditionally been seen to be tangible

and direct, although settling this question lies outside the scope of this study. She also cautioned that the subdivisions between the dimensions were somewhat arbitrary, noting 'source, intention, and time may be more accurately characterized as continua, reflecting gray areas that fall between the levels' (Kirkhart, 2000:8). The three dimensions in Kirkhart's model are examined below.

Figure 3.1 Kirkhart's Integrated Theory of Influence (2000:8)



Source

The first dimension is the source of influence and refers to the active agent of change or the starting-point of a generative process of change. Sources can be derived from either the process or the results of the evaluation. Kirkhart defined the latter within the constructs of the three traditional dimensions of instrumental, conceptual, and symbolic use. In doing so she noted that 'these three vectors of use have delimited the conceptual landscape of results-based influence' (Kirkhart, 2000:10).

Intention

The second dimension is that of the intention of influence and refers to 'the extent to which evaluation influence is purposefully directed, consciously recognized, and planfully anticipated' (Kirkhart, 2000:11). Again this dimension has two components—intended use and unintended use—with any given evaluation able to exert 'intended influence only, unintended influence only, or a mix of the two' (Kirkhart, 2000:11).

Kirkhart (2000:12–13) described a number of characteristics of intention. Intended influence might be results-based or process-based. It might also be manifest or latent. Manifest intention referred to the stated purposes of evaluation, whereas latent intention referred to the unstated or covert intentions and needed to take account of the understandings and various agendas of the clients, the evaluators, and the various stakeholder audiences. Unintended influences might also derive from either the process or the results, but 'the nature of influence, the persons or systems influenced, and the persons exerting the influence are other than desired or anticipated' (Kirkhart, 2000:13).

To illustrate an unintended influence by means of an example adapted from Kirkhart (2000:13), suppose a school director were to conduct a needs assessment on the problem of violence in schools. The intention is to involve parents and teachers with the director in identifying concerns and suggesting solutions for a safe school. However, students become involved in the evaluation and their participation alters the climate of the school.

Time

The third dimension, that of time of influence, is superficially self-explanatory and straightforward. As shown in Figure 3.1 above, Kirkhart (2000:14–17) has provided

three chronological subdivisions of the time when influence occurs: immediate, end-of-cycle, and long-term. She cautioned, however, that time is a continuum and therefore these subdivisions are arbitrary and that the subdivisions refer not only to the passage of time but also to the stages through which the processes of evaluation move. She noted that these stages are parallel to Scriven's (1991, cited in Kirkhart, 2000:14) view of program outcomes as immediate, end-of-treatment, and long-term. The three categories direct attention to three different stages and to 'any influence that is visible in a given time period, whether it is an event occurring only within that period or a process that is flowing through it' (Kirkhart, 2000:15). For example, an immediate influence could occur at the time the evaluation was taking place and then disappear or it could begin during the evaluation and continue beyond it so as to be visible in the end-of-cycle and long-term time stages. To illustrate an influence being visible over time, again adapted from an example by Kirkhart (2000:17), suppose a school establishes a focus group to collect data on the success or otherwise of an anti-bullying program. This evaluation is used to provide immediate feedback to the school, adjustments to the program are made accordingly, and the principal then decides to include the achievements of the program in the school's annual report. As a result, the regional director, impressed with the report, provides funding to expand the program to other schools in the region.

Immediate influence first occurs or is visible during the anticipating, planning, and implementing phases of the evaluation process. Immediate use therefore can be fast-paced or protracted. End-of-cycle influence commences after the evaluation process itself has concluded in a summative evaluation or at the end of a cycle in a more formative evaluation. Kirkhart (2000:16) included in this phase the influence that 'emanates from both the products of the evaluation (for example, reports, summaries, and other documents) and the process of disseminating results' in order to highlight the influence 'surrounding the conclusion of a summative evaluation study or of a cycle in a more formative evaluation'. Long-term influence looks at future use, at effects that may not be evident until well after the evaluation, or effects that are long lasting. Long-term influence, recognised by Kirkhart (2000:17) as significant, has also

been found (Cummings, R., 2002:4; Shulha & Cousins, 1997) to be often absent from empirical studies of use.

By way of caution, Kirkhart (2000:17) concluded that

> the incremental nature of influence should not be obscured by the demarcation of three time periods. The intention is to cue consideration of a full range of influence across time rather than restricting reflection to a narrow band. Hence the time dimension helps one attend to both the pace of change and the chronological periods in which it is evidenced.

Thus immediate influences merge into end-of-cycle influences, these in turn merging into long-term influences, which Kirkhart noted could date from the evaluation process itself or take years to emerge or both. This is significant because the intention of the present study is to understand the progression of influences through the immediate – end-of-cycle – long-term continuum. For example, within the Cyclical Review process, an immediate influence could be that a participating principal would gain awareness and knowledge of evaluation techniques that he had never previously used in his school. This same principal may then use these techniques for subsequent evaluations. Having gained confidence in these techniques, he may in years to come seek out other techniques and incorporate them into ongoing evaluations.

Written in 2000 Kirkhart's work preceded much of the rise in interactive approaches to evaluation—transformative participatory evaluation, practical participatory evaluation, and evaluation capacity-building. These approaches have led to three things: an increase in the transparency of data being made available to all stakeholders; stakeholder involvement—engagement, interaction, and communication—throughout the evaluation; and a closing of the gap between completion of the evaluation process and the dissemination of a report (where there is one). One effect of such an approach is the blurring of stages as processes become concurrent.

Since 2000 Kirkhart's work has been further developed by both Caracelli and Preskill (2000) and in particular by Henry and Mark (Henry & Mark, 2003; Mark & Henry, 2004) who have developed a model and integrated theory of influence (see Figure 3.4). The next section examines this model and theory.

Henry and Mark's Definition of Influence and Schematic Theory of Influence

Henry and Mark (2003) and Mark and Henry (2004) also argued that the term *use* was no longer adequate, 'advocating for moving the field of evaluation beyond use and toward a focus on evaluation influence' (Henry & Mark, 2003:295), as proposed by Kirkhart (2000). At the same time they extended the concept to include the mechanisms and outcomes of evaluation that influence attitudes and actions. They further argued that the ultimate goal of evaluation is social betterment rather than use. This definition of *influence* combines the dictionary meanings of both *use*—'the purpose for which a thing can be used'—and *influence*—'the effect a person or thing has on another' ("Australian Concise Oxford Dictionary," 1997).

The literature above highlighted a range of key gaps and shortcomings in the literature on use—for example, differentiating *use* from *misuse*—, the application of qualitatively different attributes to distinguish between the different types of use, and the narrowness of the conceptual and linguistic base (discussed in previous sections of this chapter). In addition, Henry and Mark (Henry & Mark, 2003; Mark & Henry, 2004) identified another major gap in which they asserted the literature on use was silent. This was the range of underlying mechanisms that occurred personally, interpersonally, and at an organisational or collective level, and through which evaluation might have its effects on a path towards social betterment. They drew these mechanisms from change processes within the social and behavioural sciences.

Mechanisms and processes

In 2003 Henry and Mark classified fifteen mechanisms in terms of the three levels of analysis. In 2004 Mark and Henry proposed a two-dimensional framework that not only expanded the list of mechanisms but also classified them into four distinct types: general influence processes; cognitive and affective (or attitudinal) processes; motivational processes; and behavioural processes. This classification scheme is shown in Figure 3.2. The framework was designed to show corresponding processes within each row and across each level of analysis. For example, within the general influence processes, 'elaboration at the individual level, persuasion (and the dialogue that often is involved in persuasion efforts) at the interpersonal level, and the various forms of collective deliberation, are roughly corresponding forms of information processing at different levels of analysis' (Mark & Henry, 2004:43).

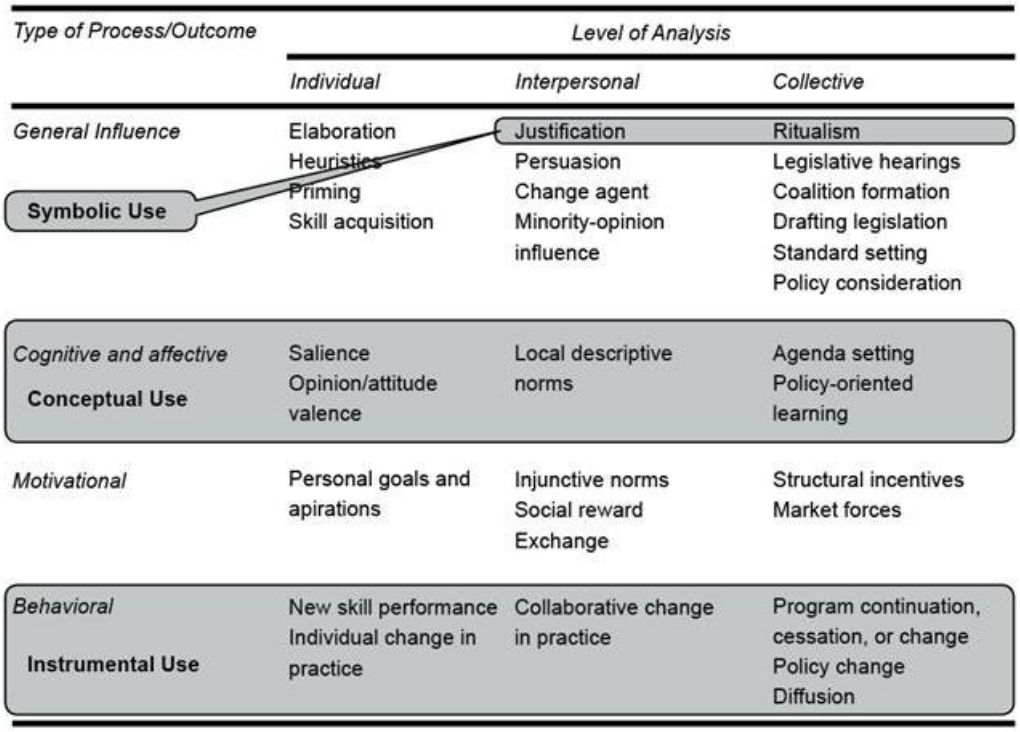
In referring to each of the entries (Figure 3.2 overleaf) as processes, Mark and Henry (2004:43) acknowledged that each played a dual role: an outcome of the evaluation or, in turn, a mechanism that gave rise to another outcome. Henry and Mark (2003:299–305) and Mark and Henry (2004:40–43) provided explanations of each of these mechanisms and their classification; these mechanisms and classifications will inform the remainder of this study, because they provide a useful classification for investigating and differentiating all forms of influence on the participating principals and the manner in which one influence leads to another.

Figure 3.2 Mark and Henry's (2004:41) Model of Alternative Mechanisms that May Mediate Evaluation Influence

Type of Process/Outcome	*Level of Analysis*		
	Individual	*Interpersonal*	*Collective*
General Influence	Elaboration Heuristics Priming Skill acquisition	Justification Persuasion Change agent Minority-opinion influence	Ritualism Legislative hearings Coalition formation Drafting legislation Standard setting Policy consideration
Cognitive and affective	Salience Opinion/attitude valence	Local descriptive norms	Agenda setting Policy-oriented learning
Motivational	Personal goals and apirations	Injunctive norms Social reward Exchange	Structural incentives Market forces
Behavioral	New skill performance Individual change in practice	Collaborative change in practice	Program continuation, cessation, or change Policy change Diffusion

Moreover, Henry and Mark (2004:43–45) asserted that their framework provided a 'more specific way of describing and operationalizing' the traditional forms of evaluation use and perceived correlations between their framework and the traditional forms of use. For example, instrumental use usually correlates with the behavioural row, conceptual use with the cognitive and affective row, and symbolic use more narrowly with the justification and ritualism processes of the general influence row. In Figure 3.3 the researcher shows her interpretation of the correlation between the traditional forms of use and those of Figure 3.2. The mapping in Figure 3.3 demonstrates that no aspect of the traditional forms of use is lost in adopting Henry and Mark's framework; on the contrary Henry and Mark's framework is an enhancement of the traditional forms of use.

Figure 3.3 Correlation between Traditional Use and Mark and Henry's Model



Type of Process/Outcome	*Level of Analysis*		
	Individual	*Interpersonal*	*Collective*
General Influence	Elaboration Heuristics Priming Skill acquisition	Justification Persuasion Change agent Minority-opinion influence	Ritualism Legislative hearings Coalition formation Drafting legislation Standard setting Policy consideration
Symbolic Use			
Cognitive and affective **Conceptual Use**	Salience Opinion/attitude valence	Local descriptive norms	Agenda setting Policy-oriented learning
Motivational	Personal goals and apirations	Injunctive norms Social reward Exchange	Structural incentives Market forces
Behavioral **Instrumental Use**	New skill performance Individual change in practice	Collaborative change in practice	Program continuation, cessation, or change Policy change Diffusion

In addition to the correlation across rows, the framework allows for clarifications and distinctions within the rows at three levels of analysis. This is important in trying to achieve a particular type of use. For example,

> if textbooks tell evaluators only how to try to achieve instrumental use, this ignores the fact that pathways to change in individual practice are likely to differ greatly from the pathways to collective change including policy change.
>
> (Weiss, 1998b, cited in Mark & Henry, 2004:44)

This aspect of the framework is potentially useful because it allows differences across alternative spheres of evaluation practice to be made. In the case of this study, the influence the Cyclical Review had on participating principals as individuals, the influence it had within their school as practitioners (interpersonal), and the influence it had on participating principals as policy-makers (collective) can be differentiated.

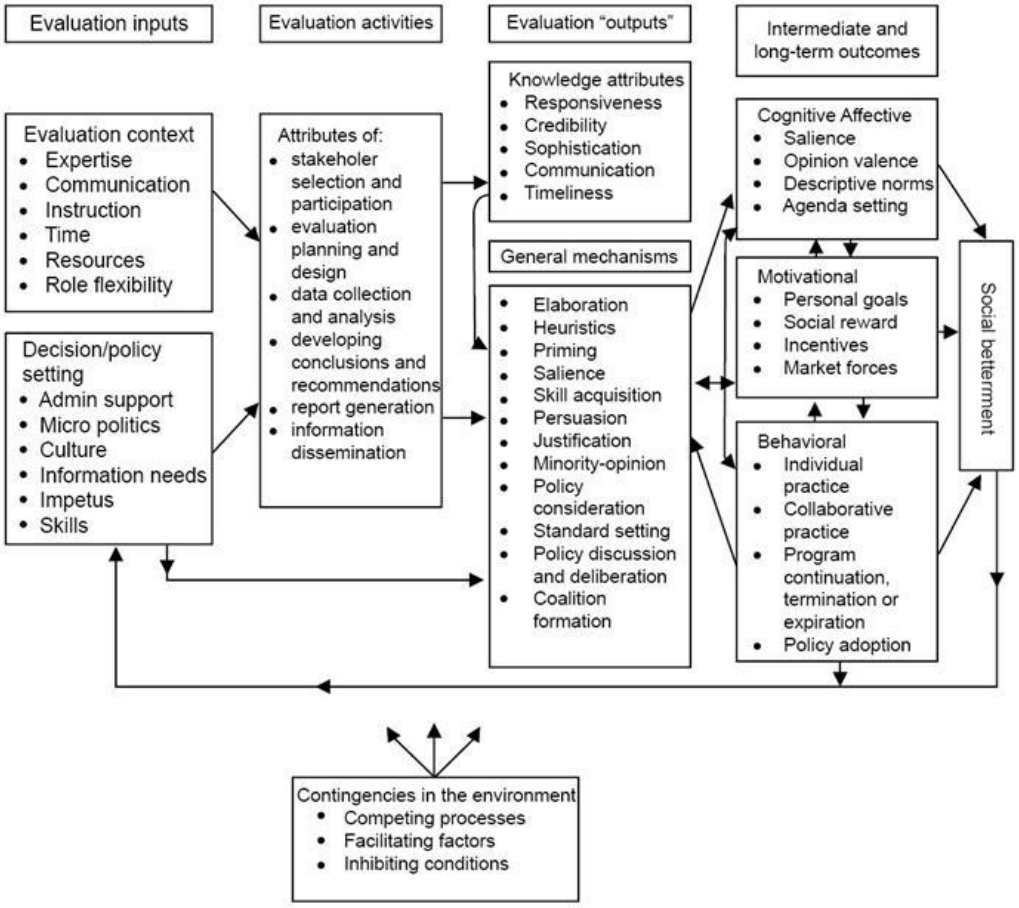
The work of Mark and Henry has moved theorising about evaluation from use and utilisation to influence and is further enhanced by their later work (Mark & Henry, 2004) on a schematic theory. In this later work, and starting from Cousins's (2003) conceptual framework for participatory evaluation, Mark and Henry (2004) embed their two-dimensional matrix into a broader schematic theory of evaluation influence.

While an enhancement in many regards, the schema itself does not show as clearly the alternative spheres of influence discussed above. Further, Mark and Henry stated that mechanisms of use referred only to results-based use, whereas the researcher contends that mechanisms can also be process-based. Mark and Henry's schematic theory and its limitations for this study will now be discussed.

Schematic theory of evaluation influence

In their schematic theory of influence (Figure 3.4) Mark and Henry (2004) suggested three main benefits over past theories of use or influence. First, the theory went beyond the factors that were precursors to or influences of use and included 'change processes through which evaluation influences attitudes, motivations and action' (Mark & Henry, 2004:45). Second, it could 'help organize a number of different hypotheses that could guide future research on evaluation influence' (Mark & Henry, 2004:45). Third, it offered a 'set of linkages between the three different levels of evaluation outcomes and the bidirectional pathways between the underlying mechanisms and the outcomes' (Mark & Henry, 2004:47). Where Kirkhart's (2000) theory allowed influence to be multidirectional, Mark and Henry's theory advanced that notion by predicting where the unidirectional and multidirectional influences occurred, as shown in the following figure.

Figure 3.4 Schematic Theory of Evaluation Influence (Mark & Henry, 2004:46)



Nevertheless, despite its utility in the context of this study, the schema presents some limitations. First, as discussed in the previous section, the schema allows for three levels of analysis but does not show these levels clearly. Second, Henry and Mark (2004:44) also contended that

> in contrast to other traditional forms of use, process use does not correspond to specific rows or processes [in Figure 3.2]. Instead, . . . process use means that some change has arisen because of the process of an evaluation rather than because of an evaluation finding. Accordingly, process use, in our view, cannot be translated into the mechanisms of [Figure 3.2]. Instead, process use is defined by whether influence is triggered, in the language of logic models, by evaluation activities rather than by evaluation outputs (i.e. findings)

The researcher's experience would lead her to contest this view, at least in the context of the present study. For example, if a participating principal acquires a new skill, such as the use of a new computer application relevant to evaluation, and then encourages his or her staff to use this application for their own work, the researcher would contend that this is process use of the evaluation and it can also be described as a behavioural process and outcome that has occurred at both the personal and interpersonal levels. Awareness of this possibility enables the researcher to recognise such occurrences and allow for them in the framework to be developed specifically for this study.

Nevertheless this schematic theory is useful to the current study because, as Mark and Henry (2004:52) asserted, it allows modifications that can be localised to a particular context. The more general framework provides for concepts to be adapted to develop more specific local theories of evaluation influence for a given evaluation.

EVALUATION UTILISATION, USE, OR INFLUENCE?

The literature discussed thus far considers four issues regarding the terms *utilisation*, *use*, and *influence*. Three of the issues are interrelated. The first is concerned with the dictionary meanings of these terms and the extent to which these meanings have been maintained in their application to evaluation. The second issue is whether evaluation use is different from evaluation influence. The third issue is, therefore, what term or terms should be used when referring to the effects of evaluation.

It has been shown above that, generically, *use* and *influence* are not synonymous, although in current usage *utilisation* and *use* are. Just as the generic meanings of *use* and *influence* are not identical, evaluation researchers for the most part have also made distinctions between the conceptual meanings of these two terms. Kirkhart's (2000:7) definition of influence included the terms *capacity* and *power*. But her definition went further, characterising evaluation influence as 'intangible or indirect' (Kirkhart, 2000:7),

compared with evaluation use. Although Kirkhart suggested that evaluation influence was 'broader than use' (p. 7), Kirkhart's definition appeared to exclude use, rather suggesting that the two aspects should sit 'alongside' each other. Nevertheless, Kirkhart argued that the term *evaluation influence* should replace *evaluation use* so as not to 'constrict our understanding of the impact of evaluation' (Kirkhart, 2000:7).

R. Cummings (2002) favoured the term *influence*, suggesting that the term implied a more deliberate but incremental change. He therefore saw the concept of influence as more active, with a greater expectation on the evaluator to effect change. It is interesting that Cronbach (1980:53) suggested a similar meaning. He and his associates suggested that 'influence comes from engagement, not detachment', therefore linking influence to more participatory forms of evaluation that have evolved in the intervening period. That *influence* may be a better term to describe participatory forms of evaluation is of relevance to this study, because, as will be shown later in this chapter, Cyclical Reviews have many of the characteristics of participatory evaluation, and indeed were developed to involve the principals as evaluators and participants.

Although not excluding the above definition, Mark and Henry (2004:40) adopted the term *evaluation influence* because they believed that 'it explicitly includes both changes that take place at the location and general time frame of the evaluation and changes that take place elsewhere and later'. This suggested that in their view *use*, both linguistically and conceptually, was a subset of *influence*, and hence *evaluation influence* was appropriate. Again it is relevant to this study to have this further concept captured in any term to be used, because the direct and indirect influences over time are seen to be necessary to understand fully the effects of Cyclical Reviews on the participating principals.

Alkin and Taut (2003), however, distinguished between the two terms and believed both had different meanings and referred to different concepts. *Evaluation use* 'refers to the way in which an evaluation and information from the evaluation impacts the

program that is being evaluated' (p. 1). *Evaluation influence* 'adds to the concept of use in instances in which an evaluation has *unaware/unintended* impacts' (p. 9). Further, they stated that 'influences of evaluation are undoubtedly of importance, but they are unintended and cannot be addressed until after they have occurred (p. 10). This is a different definition and concept of influence from those of Kirkhart (2000), restricting influence within the dimensions of intention (including only unintended influences) and time (including only long-term influences). Using these definitions, Alkin and Taut (2003) asserted that use has primacy for evaluators because, although influence might occur, it was beyond the scope of an evaluator's action. They nevertheless acknowledged the shortcomings of the traditional categories of evaluation use and suggested that *use* could be reconceptualised in terms of change mechanisms. They appeared to be arguing for the term *evaluation use*.

On the other hand Lawrenz, King, and Greensied (2005) made the distinction between the terms *evaluation use* and *evaluation influence* on the basis of purpose, and proposed that both terms be used. This distinction closely aligns to one of the differences in meaning between the terms as given in *The Australian Concise Oxford Dictionary* (1997).

Rossman and Rallis (2000:59) did not distinguish between the terms *use* and *influence* but included a learning dimension, stating: 'if we view evaluation as learning and see learning as a socially constructed, appreciative process, then evaluation use becomes reconceptualised as continual and collective knowledge generation and application'.

The problem with the various researchers' positions on terminology, as described above, is that the multiple concepts that have been identified as relevant and important to this study are not fully contained within any one of their positions. Using multiple terms and trying to distinguish between them would add another unwarranted layer of complexity to the study.

Therefore, based on this literature, it would appear that there are differences between the terms *use* and *influence* and that as yet there is no agreement on whether one or the other should take precedence. It may be that another term altogether will emerge that links the two terms and concepts together. For the purposes of the research reported here, however, the term *evaluation influence* will be adopted to refer to the intended and unintended effects over time that Cyclical Reviews have on participating principals through direct or indirect means.

An appropriate question that arises for this study, based on the literature referred to above and the definition of *evaluation influence* adopted by the researcher, is therefore: How does participation in Cyclical Reviews in Western Sydney Region influence participating principals? This research question, and others, will be revisited at the end of this chapter, before a suitable methodology is discussed in Chapter 5.

The fourth issue is the development of a framework or theory of evaluation influence (or use) that can categorise this complex concept in a meaningful way. Those of Kirkhart (2000) and Henry and Mark (2003; Mark & Henry, 2004) have been explained in previous sections of this chapter. These authors also suggested that this work is by no means complete. Mark and Henry (2004:50), for example, went on to suggest that their framework did not provide 'the ultimate, final classification and listing of the mechanisms that underlie evaluation influence'. They listed several limitations and aspects for future development, including the relative emphasis of mechanisms in particular evaluations, the modification of the framework for specific contexts, and the attention to 'various complexities that impinge on evaluation influence processes' (p. 50). Based on the work of Kirkhart (2000) and Mark and Henry (2004) and the definition of *evaluation influence* adopted for the research reported here, it became evident to the researcher that pre-existing frameworks for examining the data and addressing the research questions for this study would not suffice. The question that arises for this study, based in the literature above is whether a specific framework for Cyclical Reviews can be developed that will allow all the influences on participating

principals to be mapped. More formally, an appropriate research question is: Is there a theoretical framework that can be designed to map the influence of Cyclical Reviews on the participating principals? This research question, and others, will be revisited at the end of this chapter, before a suitable methodology is discussed in Chapter 5.

FACTORS THAT MAY TRIGGER EVALUATION INFLUENCE

The research cited thus far in this chapter on use and influence assumes that people who are influenced by the evaluation will interact in some way with the evaluation: actively or passively; directly or indirectly. The triggers that cause them to do this have been the subject of numerous empirical studies and theoretical research since the late 1970s (Johnson, K. et al., 2009). There are extensive, major, and often-cited reviews of these studies and research, including those of Leviton and Hughes (1981), King and Thompson (1983), Alkin (1985), Cousins and Leithwood (1986), Shulha and Cousins (1997), Cousins (2003), and K. Johnson et al. (2009).

Leviton and Hughes's 1981 study identified five potential factors that might lead to use of evaluation findings. They were the:

1. relevance of the evaluation to the users' needs
2. quality, quantity, and methods of communication and dissemination
3. way the information is processed
4. perceived credibility of the evaluator and the evaluation information
5. degree of user involvement and advocacy.

In their literature review of empirical research on school use of program evaluation in local educational settings, King and Thompson (1983:11–14) identified factors affecting use, their premise being that identification of critical factors may be the key to improving the use of evaluation information at the district level. They asserted that these factors can be logically subdivided into those that cannot be controlled by the

evaluator (for example, the setting) and those that can (for example, the quality of the methodology). The literature review further identified that it is the interaction between the evaluator, the users and their context, and the evaluation information that affects use. Alkin's (1985) study grouped multiple characteristics into three factors: human, context, and evaluation. In this classification, the human factor includes both evaluator and user characteristics, such as the evaluator's credibility and the user's commitment to use; the context factor looks at such characteristics as the features, requirements, and constraints of the organisation; and the evaluation factor includes the procedures, methods, and models that the evaluation adopts. These studies are of particular relevance to this study because the variables relating to human, context, and evaluation factors, and their interaction were also considered important when the Cyclical Review process was developed, as outlined in Chapter 2.

In a seminal study on evaluation use, Cousins and Leithwood (1986:347–348) identified, from sixty-five empirical studies, twelve characteristics affecting evaluation use and arranged them into two clusters. The first cluster contained characteristics related to the evaluation implementation and the second contained characteristics related to the evaluation setting, as shown in Table 3.1 below.

Table 3.1 Characteristics affecting evaluation use (Cousins & Leithwood, 1986:347–348)

Evaluation Implementation	**Evaluation Setting**
evaluation quality	commitment
credibility	information needs or competing information
relevance	personal characteristics
communication quality	decision-making
findings	political climate
timeliness	financial climate

Going beyond this table of possible factors, Cousins and Leithwood (1986:349) then developed a *Prevalence of Relationship Index*, S/R×O, where S represented the number of

studies examined, R the number of reports examined, and O the number of observed relationships within both S and R to assess the relative importance of each of the identified characteristics in predicting use. The index indicated that evaluation quality was the most important factor, followed by decision-making, personal characteristics, findings, and relevance.

While these 1980s' studies focused on factors that triggered results-based use, and are therefore an important consideration for the factors triggering results-based influences in this study, it also appeared to the researcher from her work in the field that these same factors could contribute to process-based influences. For example, if the Cyclical Review process was deemed to be of high quality by the participating principal on the review team, this might become a factor that would trigger a subsequent influence, such as the principal considering a review in his or her own school.

In 1997 Shulha and Cousins undertook a review of the literature on evaluation use, including theoretical as well as empirical research that had been published since 1986. In doing so they uncovered four main developments. First, context had become a critical consideration in understanding and explaining use. For example, Preskill (1991) proposed that factors affecting use should be interpreted within a cultural context and that to do this required an understanding of the basic assumptions, management practices, and symbolic meanings of the organisation. Second, process use had become more widely accepted and recognised as a legitimate and significant form and consequence of use. Third, the rise in collaborative or participatory modes of evaluation had expanded the concepts of use from the individual to the organisational level. Fourth, the role of the evaluator had diversified to include 'facilitator, planner, consultant, trainer, or educator' (Shulha & Cousins, 1997:204).

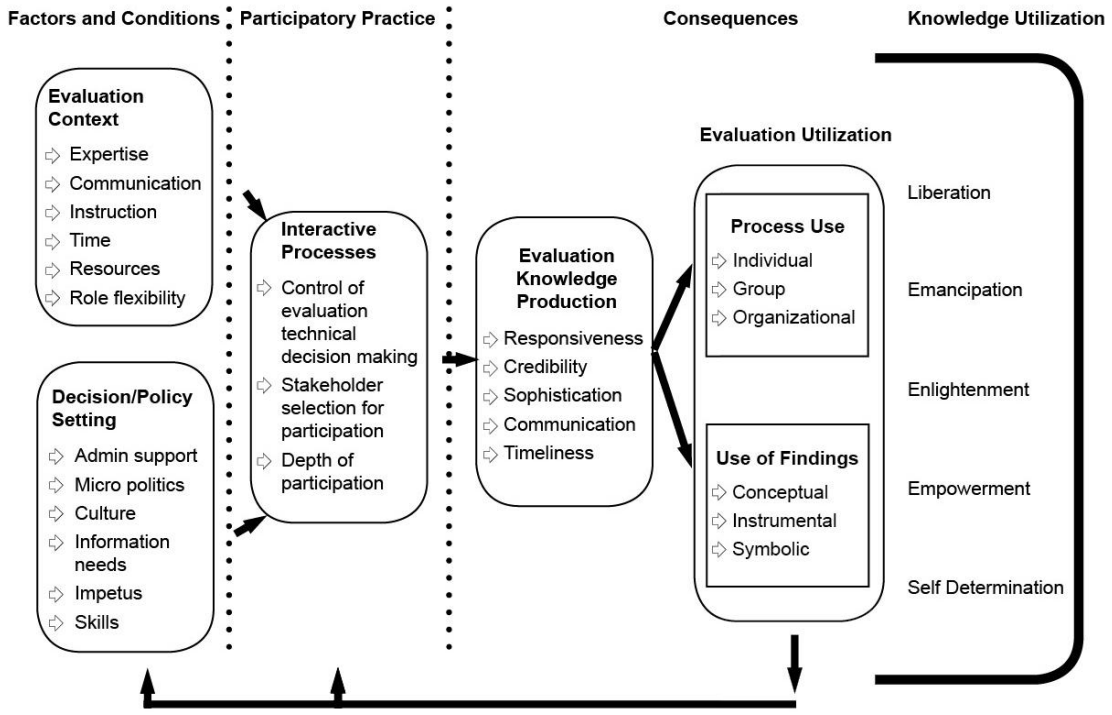
Each of these developments is considered relevant to the current study. The first suggests that a participant–observer approach is required; the second reinforces that process use is an important consideration; the third reinforces the need for a

delineation of levels of influence (that is, beyond that of an influence on the principal as an individual) in any framework to be developed specifically for Cyclical Reviews; and the fourth captures the multiple roles that the principal leading a Cyclical Review may be expected to fulfil.

In 2003 Cousins proposed a linear conceptual framework for participatory evaluation linking participatory evaluation use and knowledge use. This is shown in Figure 3.5 overleaf. In common with Cousins and Leithwood's (1986) study were many of the same individual characteristics that influenced use, grouped under the sections evaluation context, decision/policy setting, and, to a certain extent, evaluation knowledge production. What was different was the addition of participatory practices, which were also seen as influencing use, and the recognition that a combination of all these factors might also lead to process use. This framework appears to support the researcher's contention from examination of the 1980s' literature and her work in the field that factors influencing results-based use may also influence process-based use.

In addition, the identification of three interactive processes contributing to evaluation influence—control of evaluation technical decision-making, stakeholder selection for participation, and depth of participation—provided a useful starting-point to consider the range and types of interactive processes that had been built into Cyclical Reviews and the intended influences of these processes (as described in Chapter 2).

Figure 3.5 Framework for Participatory Evaluation (from Cousins, 2003:248)



Preskill, Zuckerman, and Matthews (2003) focused specifically on factors affecting process use in their study with the American Cancer Society through the Collaborative Evaluation Fellows Project (CEFP), evaluating the *Tell A Friend* program. Their study is of particular relevance to the current study because it explores what and how stakeholders in a participatory evaluation process, not unlike the Cyclical Review process described in Chapter 2, learnt from their involvement in the process. Their study identifies five categories of variables that appear to affect process use, noting that 'because it is hypothesized that process use is best supported through collaborative and participatory evaluation approaches, it should be not surprising that so many of the variables identified as contributing to process use are, more generally, related to effective group processes' (Preskill et al., 2003:430).

The five categories identified are:

1. Facilitation of the evaluation process. The twelve variables within this category relate to the extent to which a learning environment is created during the evaluation and include, for example, the 'amount of dialogue and reflection', and 'the degree of trust among group members' (Preskill et al., 2003:430).
2. Management support. The four variables within this category refer to the extent to which managers are advocates for learning, based on the rationale that 'if managers support learning from evaluation, their employers will be more likely to engage in and believe in the value of evaluation' (Preskill et al., 2003:431).
3. Advisory group (stakeholder) characteristics. The seven variables within this category that affect individual as well as group learning include, for example, 'interest in the evaluation process' and 'motivation to participate in the evaluation' (Preskill et al., 2003:432).
4. Frequency, methods, and quality of communications. The five variables within this category refer to the extent to which evaluation team members have time and opportunities to share, reflect, and stay engaged in the evaluation.
5. Organisational characteristics. The six variables within this category refer to the context within which the evaluation takes place and include, for example, the 'degree of organizational stability' and the 'extent to which the organization's culture supports ongoing learning' (Preskill et al., 2003:434).

In the most recent review of the empirical literature on evaluation use, K. Johnson et al. (2009) used Cousins and Leithwood's (1986) framework to review studies conducted between 1986 and 2005. In so doing they added to the framework one additional category—stakeholder involvement—and one new characteristic under the category of evaluation implementation—evaluator competence. In terms of the latter they suggested that it went further than the credibility characteristic, which addressed what the evaluator did, by focusing on 'who the evaluator is' and therefore 'the influential nature of the evaluator's personal competence or leadership as a means of affecting the level of evaluation use' (Johnson, K. et al., 2009:382). Stakeholder involvement, they believed, reflected the increased focus on participatory forms of evaluation and they

concluded that 'engagement, interaction, and communication between evaluation clients and evaluators is [*sic*] critical to the meaningful use of evaluation' (Johnson, K. et al., 2009:389). These additions, they believed, also took into consideration the developments in the field, such as the increased significance of process use, the expansion of use from the individual to the organisational level, and the diverse roles being adopted by evaluators.

Despite these additions, Cousins and Leithwood were unable to identify studies that would assist in progressing the concept of evaluation influence in terms of 'identifying important variables in a sequence suggestive of a pathway' (Johnson, K. et al., 2009:388), as had been suggested by Mark and Henry (2004). They suggested that 'research on influence pathways' would 'necessitate a different strategy' (Johnson, K. et al., 2009:388). Weiss et al. (2005) found that when they tried to look at influence pathways in their DARE study, they became entangled because

> Mark and Henry's work concentrates on the processes that intervene between an evaluation study and the audience that uses it to change its thoughts and actions. Our work concentrates on the collective end result. . . . We are on less sure ground trying to reconstruct individual and interpersonal processes that were reported to us some 2 to 8 years after the events. The Mark and Henry framework may be more valuable when a study begins with the framework in mind.
>
> (Weiss et al., 2005:26)

The researchers cited thus far have devised various ways of grouping the factors that trigger evaluation influence, each grouping being suitable in its context, although none has been able to trace influence pathways. It became increasingly evident to this researcher that while all of the previous groupings had some relevance for the Cyclical Review process, none was completely satisfactory. However, Alkin's groupings, although originally intended for factors leading only to results-based use, are sufficiently general to encompass the individual factors identified by the cited researchers, including those factors identified as leading to process-based use. For example, the organisational-characteristics category of Preskill et al. (2003:432) and the

six variables within it could all be grouped under Alkin's context factor. Furthermore, as Alkin's groupings also capture the three major variables that were seen as important in the development of the Cyclical Review process, described in Chapter 2, they provide a good framework for grouping any factor triggering influence, in all of its forms, on the participating principals in this study.

Being able to group factors that trigger influence is important. Before this can be done, however, the question that arises for this study, based on the abovementioned literature, is which of these factors have contributed to the influence that involvement in the Cyclical Reviews in Western Sydney Region had on the participating principals. More formally, an appropriate research question is: What factors, prior experiences, and understandings have contributed to the influence that involvement in the Cyclical Reviews in Western Sydney has had on the participating principals? This research question, and others, will be revisited at the end of this chapter, before a suitable methodology is discussed in Chapter 5.

What still remained, however, was how to incorporate factors that may trigger influence into a framework that allowed for pathways of influence to be tracked (as raised by Weiss et al., 2005, above) and also allowed for general, intermediate, and long-term outcomes of evaluation to be both final results of evaluation and factors leading to or influencing other results (Mark & Henry, 2004). Again the previously stated research question ('Is there a theoretical framework which can be designed and confidently used to map the influence of Cyclical Reviews on the participating principals?'), is relevant here.

The literature cited above examines the factors triggering influence and the influences themselves of a broad range of functions and forms of evaluation. It remains to examine how Cyclical Reviews are positioned within this general framework. The following sections will do this by analysing the functions and forms of school evaluations and then positioning Cyclical Reviews within the context of school

evaluation functions and forms. Finally, Cyclical Reviews will be placed within participatory evaluation and capacity building evaluation frameworks.

FUNCTIONS AND FORMS OF SCHOOL EVALUATION

This chapter thus far has examined the literature and research pertaining to evaluation influence in all of its forms within the definition of educational evaluation adopted by the researcher at the beginning of the chapter; that is, the 'production of knowledge based on systematic enquiry to assist decision-making about a program [including] evaluation as the judgement of worth of a program' (Owen, 2006:18). Not yet examined is literature and research pertaining to influences on evaluation participants specifically arising from school-level evaluations that fit this definition of educational evaluation and that have functions and forms in common with Cyclical Reviews. It is therefore necessary to look now at the functions and forms of school evaluation, position Cyclical Reviews within this, and then to discuss any research that directly pertains to influences on the participants who take part in these types of evaluation.

The functions of evaluation in schools are many and varied. Nevo (2006) listed five major functions: decision-making, improvement, accountability, professionalisation, and certification, any or all of which might occur within five domains or forms of practice: student evaluation, teacher evaluation, evaluating instructional materials, evaluating educational projects and programs, and evaluating schools. As stated in Chapter 1 and described in Chapter 2, Cyclical Reviews were developed as a form of school-level evaluation. Nevo states that school-level evaluations have two main functions: improvement and accountability.

> On one hand, schools might be interested in an overall review of their educational and administrative activities in order to improve their overall functioning and performance (J. MacBeath, 1999; Nevo, 1995). On the other hand, a demand for accountability may require demonstration of the merit of the school and the extent that it fulfils its goals and meet the need of its 'clients' (Carnoy, Elmore, & Sisken, 2003; Linn, 2000). These perspectives are usually

> congruent with the distinction between internal and external evaluation, although both types of evaluation can be relevant to formative and summative evaluation.
>
> (Nevo, 2006:454)

De Grauwe and Naidoo (2004:39) concurred and concluded that

> the challenge is not to choose between accountability and quality improvement, but to find the right balance between these aims, between internal and external evaluation, between the criteria set by central authorities and those set by school staff itself, between the demands of the 'public' and the needs of the professional community.

This challenge was also recognised by Hattie (2005:12–13), who pressed for school-level evaluation to be geared towards the improvement function rather than accountability alone, and Ainscow (2005:10), who urged that it was important 'to measure what we value, rather than is often the case, valuing what we can measure'.

These challenges seem to have been heeded in many systems where, since the early to mid-2000s, accountability frameworks have been developed and where the inter-related components of school planning, school self-evaluation, and school reviews (both internal and external) are featured. A review of Australian and other public-school systems undertaken by the Western Australian Department of Education and Training (2006) showed that in these systems self-assessment and reporting are likely to be annual, with a planning cycle over three to four years, and a review cycle between three and six years.

Within these accountability frameworks the form that most often reflected the dual functions of local accountability and school improvement was self-evaluation and it became a prominent feature of all public-school systems in Australia and the school systems of several other countries, including New Zealand, England, Scotland, Ireland, Canada, Hong Kong, and Singapore (Radii, 2005:4). Conversely, despite attempts to emphasise the improvement function of school evaluation, it was still often the analysis of external test results, including the publication of achievement tables, that

systems or governments used in attempts to reinforce public confidence in schools. When it came to the reviewing of schools, however, the review (Western Australia Department of Education and Training, 2006) showed that, within the accountability frameworks, this form of school-level evaluation remained the most diverse component, serving either the accountability function, the improvement function, or a combination of the two. This would tend to suggest that the challenge of establishing dual accountability and improvement functions throughout a system's accountability framework had not always been fully realised.

As described in Chapter 2, New South Wales had at this time an accountability framework that allowed for a review cycle, but none had been developed. It was within the climate described above that the Cyclical Review process (a school-level evaluation in form) was developed to address this area of the New South Wales framework. It is therefore not surprising that the Cyclical Review process was specifically developed with dual accountability and improvement functions incorporated into its design.

The Influence on Participants in School-level Evaluations

Despite the propensity for school systems to conduct school reviews (or school inspections) as a major form of evaluation (Radii, 2005), and statements by various education systems (as summarised, for example, by Radii, 2005; Western Australia Department of Education and Training, 2006) about the aims and benefits of such reviews, there appears to be a paucity of literature and research pertaining to the influence of such reviews, and indeed a virtual lack of research on long-term influences of such processes. Nevertheless, some mention should be made of three studies of school review processes, because each case relates to the influence on the principal in the school being reviewed and also because each of the reviews under study was meant to fulfil both accountability and improvement functions, although somewhat more skewed towards accountability than was the Cyclical Review process.

In the first, a review of a Victorian primary school, the principal of the school expressed some of the mixed views common when participants are asked about the impact school inspections have had on them, stating

> Our reviewer assured me that the process would be positive and I would 'get over it'. While I became more comfortable with the process, it does not mean the process was necessarily more comfortable.
>
> So, as it panned out, the thoroughness of this method of review provided the school with a circuit-breaker and the methodology required to go forward with confidence. As a school we will certainly be better, much better, for the experience . . .
>
> (Whitehead, 2005:3–4)

There is the sense of being brought to account and the prospect of criticism for either one's self or one's school being found to be deficient in one aspect or another: in this case the internal annual self-review. Then there is the more positive sense of the method of the review being a valuable experience for the school and a means to better performances. The uncomfortable experience can become a comfortable one.

Reports of a mixed impact of school inspections on those taking part are repeated in a second study, that by Macpherson (1982), who examined attitudes to and feelings about the process and events experienced by the principal and staff of a secondary school under review in rural Victoria. The principal expressed both positive and negative views about his involvement in this external review, ranging from 'hopeful, apprehensive, eager, unsure, generally confident, very pleased with general responses' to 'sometimes frustrated by events and challenges' (Macpherson, 1982:10). Macpherson believed that the pre-eminent criterion when a review was being planned and implemented should be to address the impact such a process would have on the educators involved. Such a belief is the basis of this study of the principals involved in the Cyclical Reviews of Western Sydney Region.

As this study is of evaluation influence over time and not just during the actual conduct of a Cyclical Review, it is pertinent to consider what impact the review of the Victorian secondary school had on the opinions and actions of those involved a short while after the review visit had been completed. Macpherson found (1982:14) that it was as if expectations had been raised high in many different ways by the visit but then, as staff members expressed it, there was a feeling of void when nothing seemed to eventuate. The review had taken place but staff were left wondering whether the ordeal and conclusions warranted the preparation.

Countering these comments, more positive remarks came from the principal, who asserted that the impact of the review had brought positive benefits to the school and that the review had brought the school together in unity and purpose as never before. A big factor was the professional improvement and improved outlook demonstrated by many. While the process had been very demanding, challenging, and wearing it had been, none the less, worthwhile (Macpherson, 1982:16).

In the third study, an Ofsted inspection in the United Kingdom, Perryman (2005) examined the impact on the head and staff of a school whose performance was declared inadequate and as a result was subject to a special-measures order. Linking her argument to Foucault's (1977:170) work on discipline, Perryman argued that the use of disciplinary powers, such as special measures, is part of an increasing culture of accountability that has created a system in which disciplinary mechanisms are used widely. It is argued that by the use of 'tick lists' (Perryman, 2005:284), which define the characteristics of effective schools (Sammons, Hillman, & Mortimore, 1995) to control the language of school inspections, Ofsted had effectively defined the discourse through which schools are normalised.

The case-study school of Perryman's research (Northgate) found itself at the centre of a discourse of failure, and paradoxically the adoption of Ofsted's discourse of

improvement proved to be a key strategy for getting out of Special Measures (Perryman, 2005:284).

> Even given that the general consensus at Northgate was that inspections were 'demoralising and exhausting', as the school 'improved', teachers were able to welcome some of the effects of being in Special Measures, often in the framework of an 'improvement' discourse. They mentioned the importance of the focus of inspections, the push for improvement and monitoring improvement, and the development of good practice. At the same time, as teachers became practised in the use of the improvement discourse, so they were able to demonstrate improvement to the inspectors. Thus the discourse change could be said to have 'worked' for them.
>
> (Perryman, 2005:285)

The question here for the study of the Cyclical Reviews in Western Sydney Region is whether those involved become so familiar with the discourse of the *Exemplary Practice Statements* that they begin to parrot them in order to be seen to be committed to the process.

Whereas the teachers at Northgate reported that the school had become dictatorial with a top-down management style that disempowered staff and engendered a lack of ownership (Perryman, 2005:286), Head Teachers in the UK were often supportive of the special-measures process. Their replies suggested that it had strengthened their own position in the school and allowed them to take actions which might not have been possible under other circumstances.

While only a small number of studies could be found relating to processes that had similarities to the Cyclical Reviews, it could tentatively be suggested that within these processes those in leadership positions, and therefore more likely to be directly involved with the reviews and more aware of their purpose and process, were more likely to see reviews as beneficial. Alternatively it seems as if those less directly involved in reviews, whether teachers or parents, and who were less aware of their purpose and process, tended to develop negative views of reviews and a lesser commitment to the process and outcomes.

These messages arising from studies of the impact of involvement in inspections need to be compared with the data on evaluation influence gathered from the principals involved in the Cyclical Reviews of Western Sydney Region. The caution should be to acknowledge that the principals, like the school leaders in Victoria and the UK, have a far greater involvement and understanding of the review process they are helping to lead than those who are the subject to the Cyclical Reviews. The data may be skewed accordingly.

At the same time as public-school systems were addressing a dual accountability and improvement function in their accountability frameworks—that is, from the late 1990s until the mid-2000s (as discussed earlier in this chapter)—there was continued and growing interest in collaborative methods of evaluation in general and a concomitant interest by researchers on what and how stakeholders learn from their involvement in the process of evaluation. These developments were not lost on the Western Sydney Regional Director, who had commissioned the development of Cyclical Reviews and who had specifically required a process that not only fully involved the region's principals in all aspects of the process but also provided professional development for them in terms of increasing their evaluation knowledge and ability.

CYCLICAL REVIEWS AS PARTICIPATORY EVALUATION

The Regional Director appeared to outline many of the features of participatory evaluation when, in addressing a meeting of Western Sydney Principals (L. Wasson, 28 August 2007), he said

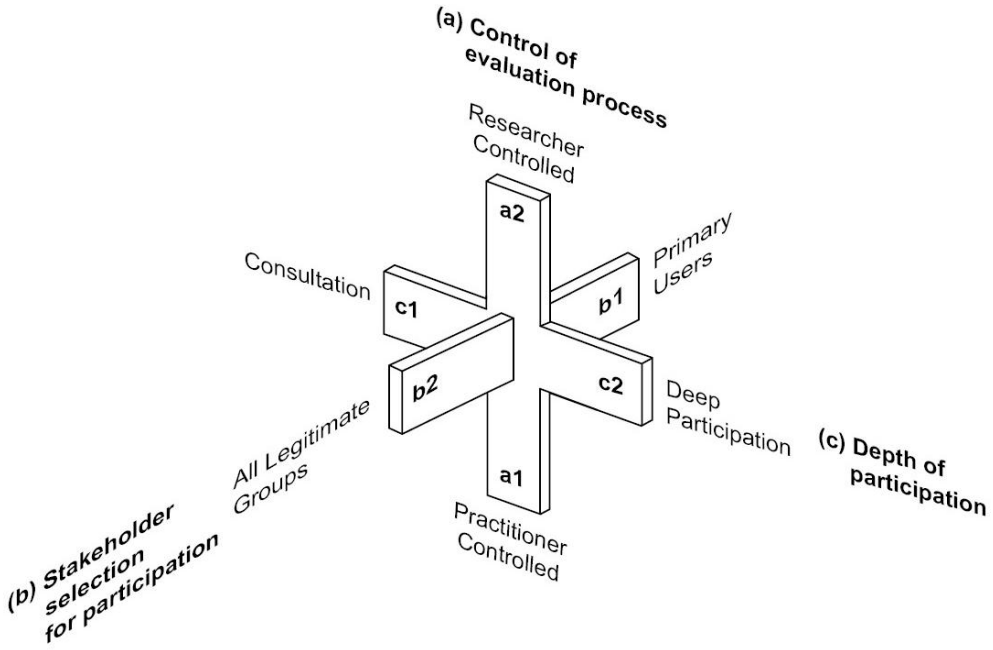
> It is anticipated that by participating in all aspects of Cyclical Reviews, principals will develop a sense of ownership of the process, further build the evaluation culture within their schools, and lead their schools to continually grow and learn.

As Kirkhart (2000:12) commented, there may be an 'explicit intention to influence organizations and social systems via the evaluation process itself, as illustrated by participatory evaluation'.

The term *participatory evaluation* has come to mean many things to many people (Cousins, 2003; Cousins & Whitmore, 2007; King, 2007), encompassing many forms of evaluative inquiry. Cousins and Whitmore (2007) defined two principal types of participatory evaluation: practical-participatory evaluation (P-PE) and transformative-participatory evaluation (T-PE). The former was focused on program decision-making and problem-solving with a central function of fostering evaluation use, the premise being that 'stakeholder participation in evaluation will enhance evaluation relevance, ownership, and thus utilization' (Cousins & Whitmore, 2007:88). The latter was focused on enlightening members of disadvantaged groups for the purposes of empowerment, the premise being that less-powerful stakeholders are empowered by participating in the 'process of constructing and respecting their own knowledge' and empowered 'through their understanding of the connections among knowledge, power, and control' (Cousins & Whitmore, 2007:90).

To differentiate between these two forms of participatory evaluation and other forms of evaluative inquiry, Cousins and Whitmore (2007) presented a framework based on the intersection of three dimensions: control of the evaluation process; stakeholder selection for participation; and the depth of participation (Figure 3.6). On these dimensions, T-PE is a1, b2, c2—practitioner control of deep participation by all legitimate groups; P-PE is a1–2, b1, c2—researcher and/or practitioner control of deep participation by primary users. The greatest difference between the two is that of stakeholder selection for participation. Despite these differences, there are also substantial similarities; for example, P-PE may empower participants, T-PE can have practical value, and both forms emphasise the importance of creating data from the practitioners' perspective (King, 2007:84).

Figure 3.6 Dimensions of Form in Collaborative Inquiry (Cousins and Whitmore 2007:93)



The parallels between P-PE and the way the Cyclical Review process was designed, as described in Chapter 2, are evident. Cousins and Whitmore (2007) stated that there needed to be an initial impetus or trigger to initiate P-PE. Once this has occurred, the two groups are developed: evaluators and practitioners. Each group needs to exhibit or be able to develop specific characteristics: technical evaluation skills; interpersonal skills; commitment to the evaluation; and the ability to work in partnership and be involved in the decision-making (Cousins & Whitmore, 2007; Greene, 1988; Preskill et al., 2003). As a result of this partnership, practitioners develop over time the ability to think and act evaluatively (MacLellan-Wright, Patten, dela Cruz, & Flaherty, 2007). The knowledge that is constructed over time reinforces the use of evaluation findings (Torres & Preskill, 2001).

Nevertheless, learning is not the primary function of P-PE. Rather the P-PE approach addresses practical problem-solving to inform decision-making (Cousins, 2003). In contrast, the Cyclical Review process was developed to fulfil dual primary functions:

practical problem-solving to inform decision-making, and development of the evaluation capacity of the region in general and principals in particular.

In terms of the first purpose, Cyclical Reviews appear to be encapsulated in Cousins's (2003:245) definition of participatory evaluations as

> an approach where persons trained in evaluation methods and logic work in collaboration with those not so trained to implement evaluation activities. That is members of the evaluation community and other stakeholder groups relative to the evaluand each participate in some or all of the shaping and/or the technical activities required to produce evaluation knowledge leading to judgements of merit and worth and support for program decision making . . . PE is distinguishable from other forms of collaborative inquiry such as stakeholder-based evaluation and empowerment evaluation by virtue of the requirement that members of both the evaluation community *and* other stakeholder groups are *directly* involved in the production of evaluation knowledge.

The significance therefore of Cousins and Whitmore's (2007) framework to this study is that it provides a useful perspective that could be used to position the Cyclical Review process within the broad and complex field of participatory evaluation. In so doing, it could more clearly define the role the principals are expected to assume, the enabling factors and issues that would affect influence (previously outlined in the section *Factors that May Trigger Evaluation Influence*), and the influences on principals that would be expected to occur at a personal level, interpersonally (as the leader in their school) and collectively (as a senior officer in the region and the system).

Owen (2006:220–221) believed that a participatory evaluation approach could be linked to ECB. Again, his description of this link shows many parallels to the Cyclical Review process outlined in Chapter 2. He saw organisations learning through a process of inquiry into practice and building an evaluation culture through an internal evaluation system. 'The existence of an evaluation culture would in turn, sustain the evaluation

component of an organisational commitment to learning' (Owen, 2006:221). ECB will therefore be discussed in the next section.

CYCLICAL REVIEWS AS ECB

The link between a participatory evaluation approach to program evaluation, in this case Cyclical Reviews, and a process that leads to ECB also appears to have been recognised by the Regional Director of Western Sydney, when at the 28 August 2007 meeting of Western Sydney Principals he further commented that

> It is anticipated that Cyclical Reviews will provide schools with valuable data and advice as to the strengths and areas for development in their teaching and learning programs, leadership, management, and culture, and will serve to build the evaluation capacity of those directly involved and in particular that of the school principals.

According to Alkin and Christie (2004), ECB arises from evaluation use, a concept discussed above. Program evaluation and its use in organisations to achieve improvement and accountability, depending on the characteristics of the process adopted, can be a useful means of building the evaluation capacity of those involved.

Thus ECB has been defined as 'the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine' (Baizerman, Compton, & Stockdill, 2002a:1). The characteristics of the evaluation process that may affect use and ECB are discussed in the literature (Baizerman, Compton, & Stockdill, 2002b; Bourgeios & Cousins, 2013; Compton & Baizerman, 2007; King & Volkov, 2005; Taut, 2007) and summarised here.

Organisational learning is a theory about how organisations learn and adapt to be more effective and efficient. The theory originated in the work of Argyris and Schön

(1978) who used the concepts of single-loop learning (that is, repeated attempts at the same problem, with no variation of method and no questioning of the goal) and double-loop learning (that is, modification or even rejection of the goal based on experience) to explain how individuals, groups, and whole organisations move from simply responding to differences between expected and actual outcomes to make improvements, to examining the deep values, assumptions, and policies that create actual outcomes and alter them.

Organisational learning has been discussed and updated over time by Argyris and Schön themselves (for example, Argyris, 1996; Argyris & Schön, 1996). In more recent years Imants (2003) has argued that two mechanisms are essential for effective organisational learning in schools; information systems about teaching and learning, and collaborative planning in professional learning communities. Another more recent variant of learning organisation theory stresses capacity building in schools through leaders taking ownership for developmental processes (Senge, 2010; Senge et al., 2000; Senge, Smith, Kruschwitz, Laur, & Schley, 2010). Jeris's (2003) study, where two organisational problem situations were worked on by sixty-one teams in four conditions—controls, traditional processes (goal clarity, commitment, and decision-making), team learning (assumption surfacing, dialogue, and reflection), and a combination of traditional processes and team learning—found that team learning produced significantly more double-loop learning than the other three conditions studied.

Another branch of organisation theory is concerned with organisational development (OD). It was founded by Kurt Lewin, who also initiated force-field analysis, action research, leadership climates, and group dynamics. Today OD is used to mount comprehensive interventions intended to boost an organisation's efficiency and effectiveness (Cummings, T., 2008). Three key strategies in OD are the enhancement of the methods used to sense and learn from changes in the internal and external

environment, the use of evaluation to provide feedback for the planning processes, and the boosting of the levels of trust between organisational members.

An even more radical and very recent suite of theoretical propositions have been termed *U learning* by Senge, Scharmer, Jaworski, and Flowers (2008), Scharmer (2009), and Scharmer and Kaufer (2013). By this they mean using a combination of double-loop learning, deep learning, organisational learning, and OD concepts and processes to radically reconceptualise and reconstruct organisations to represent the intersection of complex internal and external eco-systems.

Preskill and Torres (1999) and Russ-Eft and Preskill (2001) suggested that organisations which emphasise learning develop an improved ability to use evaluation data, an improvement that ECB seeks to achieve. In the process, skills and understandings in evaluation are enhanced as the evaluators collaboratively share, communicate, and apply data in an endeavour to improve the effectiveness of the programs (Patton, 1997; Preskill & Torres, 1999).

OD is also seen as a manifestation of ECB. The concept centres on organisations improving their capacity to address variations of context and improve effectiveness by way of a planned and continuous process (King & Volkov, 2005). If evaluation is to be linked to OD in this way, then ECB, according to Mott (2003), would include cooperative planning and systematic interventions to integrate evaluation processes into all functions of the organisation. Considering the overall context of the Cyclical Reviews, there would need to be appropriate interventions for ECB to give rise to OD.

Evaluation capacity can also be theorised as an ability to revise organisational structures, or in Giddens's (1984) terms, to manage structuration. His theory relied on an assumption about the duality of structure, by which he meant that organisational structures were both process and outcome; the former enabled changes to structure to

be imagined and created, while the latter was the both the consequence and the basis for more structuration. In total, he conceived of organisational structure as a temporary and socially constructed artefact that could be improved by the evaluation of deep values and assumptions about the best way to be organised.

ECB can therefore be viewed as a process for enhancing the skills and understandings of those charged with conducting program evaluations that then prove to be more useful to an organisation than they would have been otherwise. As such it is anticipated that the skills to plan, design, collect, and analyse data and to interpret and report results will be developed within those involved (Boyle, Lemaire, & Rist, 1999; McCoy, Rose, & Connolly, 2013; McDonald, Rogers, & Kefford, 2003).

Stufflebeam (2002) also saw ECB as a means of embedding evaluation within an organisation. To sustain an evaluation system a school district should not only increase the individual's skills and understandings about evaluation, but also seek to develop positive attitudes and commitment to the value of effective evaluation programs.

In *New Directions for Evaluation*, Baizerman, Compton, and Stockdill (2002b) discussed the characteristics of ECB. They and other writers (Dabelstein, 2003; McDonald et al., 2003) identified five characteristics of ECB. The characteristics were:

1. sustaining an evaluation system
2. considering the context and needs
3. developing an overall supportive process
4. building a cooperative partnership
5. increasing evaluation use.

Milstein et al. (2002) argued that the development of evaluation capacity must always be relative to specific organisations, since they are likely to have different objectives, policies, and programs. Given that variations in resources can also enhance or restrict

ECB, it becomes clear that contextual factors within organisations need to be taken into account and adaptations made (King & Volkov, 2005). Adaptations were certainly applied to take account of contextual factors for the Cyclical Reviews.

Again, a variation in factors can enhance or hinder the support mechanisms that have been designed to institutionalise evaluation within an organisation. The factors that could affect the support given to the evaluation process include leadership, policy, structure, existing evaluation practices, and economic and political situations. King and Volkov (2005) and McDonald et al. (2003) asserted that interventions that address only some of the more active among these factors will be deficient in supporting evaluation systems. The focus should be to work with the holistic organisational factors.

ECB also aims at building a cooperative partnership among those involved (Dabelstein, 2003; King & Volkov, 2005). In the case of the Cyclical Reviews they would include the participating principals for each review and the other team members, the host principal and his or her staff, and the members of the Cyclical Review Steering Committee. Involving representative committee members in the planning and implementation of evaluation programs can develop skills and understandings about evaluation techniques and thereby the growth of ECB (Dabelstein, 2003; Stufflebeam, 2002).

Finally, the literature asserts that ECB should result in increased and more effective evaluation use (Compton, Glover-Kudon, Smith, & Avert, 2002; Gilliam et al., 2003; Stufflebeam, 2002). These researchers have argued that an increase in the use of evaluation is a major manifestation of ECB and thereby strengthens the capacity of organisations to use evaluation effectively. Such an outcome was anticipated in the quotation attributed to the Regional Director at the beginning of this section. When members of organisations become aware of the benefits and more confidently make use of evaluation techniques, then ECB is more likely to be sustained (Cousins, Goh, Clark, & Lee, 2004).

Like most change mechanisms, ECB can be tracked through three broad stages: initiation, implementation, and institutionalisation (Fullan, 2001). As was the case with the Cyclical Reviews, ECB is usually initiated by a change agent (in this case the Western Sydney Regional Director), who is charged with increasing evaluation use in a district (region) but finds that additional support is needed from the staff (school education director, principals) and the community (schools to be reviewed).

To implement ECB an advisory committee is often established (as was the case for the Cyclical Reviews) to assess the context of the evaluation program including the factors that may or not facilitate it, develop structures and procedures, and identify resources needed to carry out the evaluation program. In doing so the committee may use surveys or interviews or scrutinise documents (King & Volkov, 2005).

The institutionalisation of ECB can only be achieved by the rigorous monitoring and evaluation of its progress. The evaluation of ECB, according to Ingram et al. (2002) and King (2002) could include a consideration of inputs, outputs, outcomes, and impact. The advisory committee could be charged with the task of planning this ongoing evaluation and in the process enhance their own evaluation skills and understandings.

Given the congruence between the aims, features, and beneficial outcomes of ECB found in the literature above, the intent of the Regional Director in commissioning the Cyclical Reviews as quoted above, and the features of the Cyclical Review process described in Chapter 2, the final research question to emerge is: To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews? This research question, with others that emerged throughout this chapter, will now be revisited, before a suitable methodology is discussed in Chapter 5.

SYNTHESIS OF KEY QUESTIONS FOR RESEARCH

The literature chosen for this review was expected to address the three broad purposes of this study, as stated in Chapter 1, namely:

1. To identify and interpret how the eighteen principals who had participated in the pilot Cyclical Review program in Western Sydney Region had been influenced over time by both the review process and review findings.
2. To gain insights into how, when, and to what extent Cyclical Reviews shape, effect, support, and change the views and practices of the principals who participate in them.
3. To use the knowledge gained from 1 and 2 above to modify and strengthen the Cyclical Review program in Western Sydney Region and, in so doing, contribute to the statewide implementation of that element of the New South Wales *School Development and Accountability Framework*.

The previous sections of this chapter presented an analysis of the range of literature relevant to the study, including that relating to accountability and development, the inspectorial system, and quality assurance, all of which fall within the broad topic of evaluation. In particular, a theoretical background to the dual concepts of evaluation use and evaluation influence was provided. In so doing, the theoretical models suggested by researchers as a means to record and analyse data relating to these dual concepts were discussed.

This section identifies the major recurring themes that emerged collectively from an examination of the literature presented in this chapter, and the context of the study in previous chapters. These give rise to the research questions for this study. These themes underpin the research questions and provide a frame of reference for the methodology, which, in turn, frames the data-gathering from the participants and the analysis in this study.

Theme 1: A Model to Record and Analyse the Data in order to Determine Evaluation Influence

It is one thing to decide that the recording of the effects on the participating principals of involvement in Cyclical Reviews is a valuable research project and that evaluation influence is the concept that best embraces all of the likely effects; it is another to identify or create a theoretical model that will facilitate such recording.

This challenge has not been ignored in the literature. Various models for measuring evaluation use and evaluation influence have been proposed. All have their basis in various researchers' definitions of one or both of the concepts and each has its own rationale and characteristics. In particular, the models proposed by Kirkhart (2000) and Mark and Henry (Henry & Mark, 2003; 2004), although not entirely applicable to this study, were considered to have many features that should be included in a new model devised for the purpose.

The discussion of the theoretical models proposed in the literature and the conclusion that a new model would be needed to undertake this study led directly to the first research question:

1. Is there a theoretical model that can be designed to map evaluation influence?

Theme 2: Knowledge, Prior Experiences, and Factors that Influence Principals who Participate in Evaluation Processes

It stands to reason that if the third purpose of the study (that is, to modify and strengthen the Cyclical Review process) as stated above, is to be achieved, then the initial factors that trigger any subsequent influence need to be determined.

The majority of the scholarly work reviewed, both in the context of the study and the literature review, sought to identify the factors that contribute to effective evaluation for accountability or developmental purposes or a combination of both. The purpose of this was twofold. From the perspective of the literature review in particular, the purpose was to identify a large range of factors that trigger influence from evaluations in general and also to examine various categorisations to group these factors. From the context of the study in particular, the purpose was to identify specific factors contributing to influence in evaluations that participating principals might reasonably have been expected to have knowledge of or experience in—inspections, quality assurance reviews, evaluation practices in other states and territories.

Thus the second research question was:

2. What factors, prior experiences, and understandings contribute to the influence that the involvement in Cyclical Reviews in Western Sydney Region has had on the participating principals?

Theme 3: The Distinction between Evaluation Use and Evaluation Influence

Again, if the first and second broad purposes of the study as stated at the beginning of this section were to be pursued, it would clearly be necessary to consult literature and research exploring the effects that involvement in evaluation programs and processes had on participants. The earlier studies reviewed focused on the use participants made of their involvement in evaluation programs and processes, whereas the later studies saw evaluation use as a limited way to judge the effects, preferring the more comprehensive concept of evaluation influence. The literature chapter traced the evolution of this debate and in so doing drew distinctions between the definitions of the two concepts, the conclusion being that *evaluation influence* was the more appropriate term for this study as it subsumes the less useful term of *evaluation use*

(Alkin et al., 1979; Henry & Mark, 2003; Johnson, R., 1998; Kirkhart, 2000; Mark & Henry, 2004; Owen, 2006).

It followed that the third research question was therefore:

3. How does participation in Cyclical Reviews in Western Sydney Region influence participating principals?

Theme 4: Cyclical Reviews as Evaluation Capacity Building

The literature review analysed the theoretical grounds, definitions, and processes of ECB. Here the links between self-evaluation, OD, and organisational learning were explored. ECB was seen as enhancing and sustaining the ability of, in this case-study, a group of schools to conduct evaluations to improve programs and promote student learning. In the process ECB is expected to improve the skills and understandings of evaluation methods and processes in those conducting the evaluations. Taking into account the internal and external contexts that influence change, ECB can be directed towards professional development, supportive infrastructure, and access to resources.

Given the intentions of the Western Sydney Regional Director in commissioning the development of the Cyclical Review process and the congruence between his intentions and the literature pertaining to ECB, it follows that the fourth and final research question was:

4. To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?

The next chapter describes more fully the models proposed by Kirkhart (2000) and Mark and Henry (2004)—their applicability to the present study and their shortfalls—before proposing a provisional theoretical model that will be used to map and analyse evaluation influence. This framework will serve as the basis for organising the

collection and analysis of data within the qualitative methodology described in Chapter 5.

CHAPTER 4
TOWARDS A PROVISIONAL THEORETICAL MODEL FOR MAPPING EVALUATION INFLUENCE

INTRODUCTION

This chapter describes more fully the models proposed by Kirkhart (2000) and Mark and Henry (2004) for measuring evaluation influence, and identifies the characteristics within them considered applicable to the present study. It then identifies what are considered to be the shortfalls of these models if they applied to the present study. Modifications or extensions to these models made by the researcher in working towards a new theoretical model specifically devised for this study, and which will be used to address the first research question, are then explained and illustrated.

APPLICABILITY AND SHORTFALLS OF KIRKHART'S MODEL TO THE PRESENT STUDY

From the description provided in Chapter 2 of the Cyclical Review process it was expected that influence could emanate from each of the dimensions identified in Kirkhart's framework (see Figure 3.1)—source, intention, and time—, and any model specifically designed to analyse the influence of Cyclical Reviews would need to include these three dimensions.

In terms of source, the traditional literature (for example, Alkin, 2004b; Patton, 1978), as discussed in the previous chapter, tends to limit the users of results to decision-makers who are the receivers of the evaluation. In Cyclical Reviews this equates to the host-school principal. It is also possible that the team members, including the team-leader and team-member principals will also make direct or indirect use of the results or be

influenced by the results of the review. It is equally possible for the host-school principal as well as the team-leader and team-member principal to be influenced by the review process. Kirkhart's framework allows these sources of influences to be identified.

The importance of intention in Kirkhart's framework to the present study is that the intended influences derived from participation in a Cyclical Review were explicitly developed and stated and the process itself was purposefully planned with principal involvement to deliver these influences, as discussed in Chapter 2. In addition, the intended influences were not only process-based and results-based but also included areas that fell between the two. In order to develop and strengthen the Cyclical Review process, it is important to distinguish between intended and unintended influences to gauge whether the intended influences occurred, when they occurred, and the factors that led to their occurrence, as well as to uncover the same information about any unintended influences—positive or negative. From the outset it was hoped that the study would reveal other unforeseen positive influences.

As discussed in Chapter 3, one effect of interactive approaches to evaluation has been a blurring of the stages as the processes become more concurrent. The Cyclical Review process is no exception. But to develop a more complete picture of influence it is necessary to uncover effects at different stages of the process. Kirkhart's concept of a time dimension was therefore important to this study. Some modifications to the actual end-points, however, needed to be made to account for the specific nature of the Cyclical Review process.

In the case of Cyclical Reviews, the duration of the evaluation process is only three days, with the first two days primarily spent gathering and synthesising the data, but also providing opportunities for the in-school stakeholders to view these data, and reporting interim findings. The third day is again spent synthesising the data, in consultation with the host principal, but the majority of time is spent writing the report

and presenting the report to the school. Adhering strictly to Kirkhart's time dimension would place most of the processes of days one and two in the immediate level of the time dimension and most of the processes of day three in the end-of-cycle level of the time dimension. With stakeholder involvement recognised as a new category leading to evaluation use (Johnson, K. et al., 2009), it seems more appropriate to place the division between the immediate and end-of-cycle stages for Cyclical Reviews at the conclusion of day three of the review.

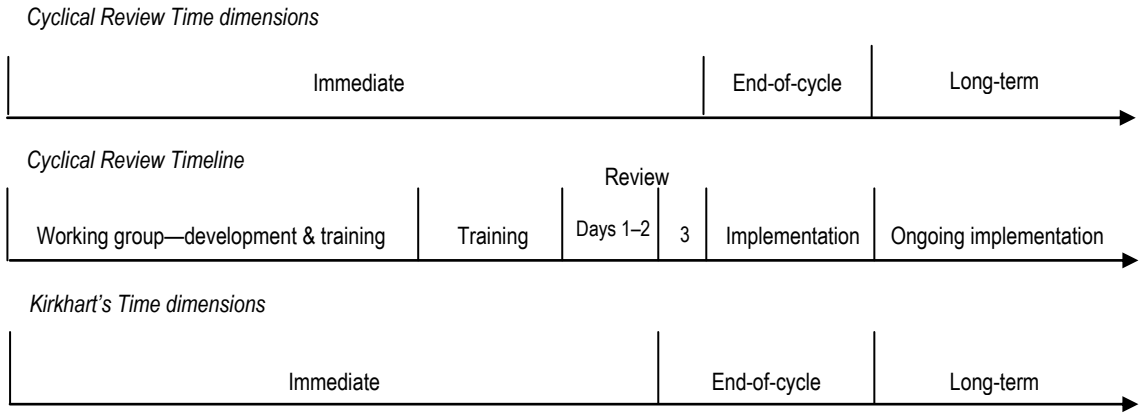
Using this division, the end-of-cycle influence, in the case of Cyclical Reviews, is seen as the processes immediately following the review's conclusion, commencing immediately after the review-team's dissemination of the report to the school under review until a point when stakeholders have had time to reflect on the products, processes, and results, disseminate them more widely, incorporate the results into plans, and begin to put plans resulting from the review into action. While there is no fixed time for these processes, they would generally be expected to occur within one to two school terms (within three to six months) following a review. The end-of-cycle influence, thus defined, is of importance to this study, because this has traditionally been the only timeframe within which influence was anecdotally seen to occur in any form of school review.

End-of-cycle influences merge into long-term influences. In this study the future time span is necessarily limited to approximately two years following the first pilot Cyclical Review. As noted by Kirkhart (2000:17), however, long-term influences could emerge years after the evaluation itself.

Where to start the immediate stage for Cyclical Reviews is more straightforward. The planning for Cyclical Reviews as a regional process began in 2006, well before the first pilot review in 2007 and the last in 2008. As the host-school principals and team-leaders also took part in planning and developing the regional process, the actual time span of immediate influence for these principals ranged from a year and a half to two

years. Principals who were team members, however, were only involved in the planning for the specific review in which they were to participate. This occurred once they were nominated, and so for these nine principals the time span for immediate influence ranged from one month to eight months. Thus in the case of Cyclical Reviews, the anticipating and planning phases are quite protracted, especially in the case of the team-leader and host-school principals, compared with the actual evaluation process for each school. A diagram showing the three periods in which evaluation influence emerges for Cyclical Reviews compared with the three periods in Kirkhart's framework is shown in Figure 4.1.

Figure 4.1 Time dimension: Cyclical Reviews compared with Kirkhart (2000)



APPLICABILITY AND SHORTFALLS OF MARK AND HENRY'S MODEL TO THE PRESENT STUDY

Mark and Henry's (2004) two-dimensional framework classifying three levels of analysis (personal, interpersonal, and collective) against four kinds of mechanisms (general influence processes, cognitive and affective, or attitudinal, processes, motivational processes, and behavioural processes) was also an important consideration for this study.

The three levels of analysis that are important to this study are: (1) individual—the principals as individuals; (2) school—the principals as leaders in their own school; and (3) system—the principals as leaders across a group of schools, at a regional level, a system level, or all three. While there are many similarities to the levels of Mark and Henry, there is an important distinction. The focus of Mark and Henry's three levels is the organisation and the influence that the evaluation has on individuals, groups, and the organisation-level decision-making within it. The focus of the levels in this study is the individual—the principal—and the influence participation in a Cyclical Review has on that individual personally, as he or she interacts with others to lead a school and implement regional and state policy and programs, and as he or she interacts with others to set or shape regional and state policy and programs.

The implication of this is that the mechanisms ascribed by Mark and Henry to one of their three levels of analysis may be ascribed to a different level of analysis in the case of Cyclical Reviews. For example, a Cyclical Review in another school may influence a team-member principal to conduct a Cyclical Review in his or her own school, resulting in a change in the latter school's policy regarding how its programs will be evaluated. The same principal may also contribute to regional discussion about Cyclical Reviews, leading to some major reforms in the review process or system-level debate and resulting in a policy change. In this case the mechanism of 'policy change' (Henry & Mark, 2003:304) could be seen to have occurred at both the school and system level of analysis in relation to Cyclical Reviews.

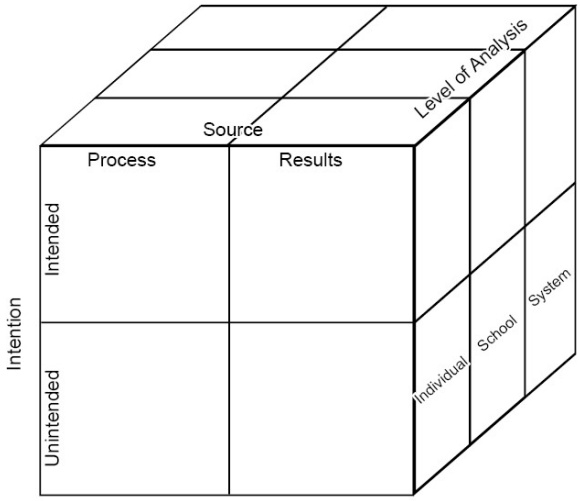
While Mark and Henry's model advances the model proposed by Kirkhart (2000), it also loses some of the distinctions proposed by Kirkhart's three dimensions, which are seen as important to this study. Mark and Henry's model does not address specifically the source—whether the mechanisms emanate from the use of the results or involvement in the process. Nor does it distinguish between intermediate and long-term influences. In particular, it does not provide a way to determine whether the influence is intended or unintended. Their model also places restrictions on some of

the pathways. At least one general influence must occur as an immediate influence and before any other mechanism or outcome. A number of pathways are unidirectional; for example, the pathways from evaluation inputs to evaluation activities, and on to evaluation outputs. Further, there is no direct feedback pathway from a general influence to a trigger for evaluation.

PROVISIONAL THEORETICAL MODEL FOR ANAYLSIS OF EVALUATION INFLUENCE

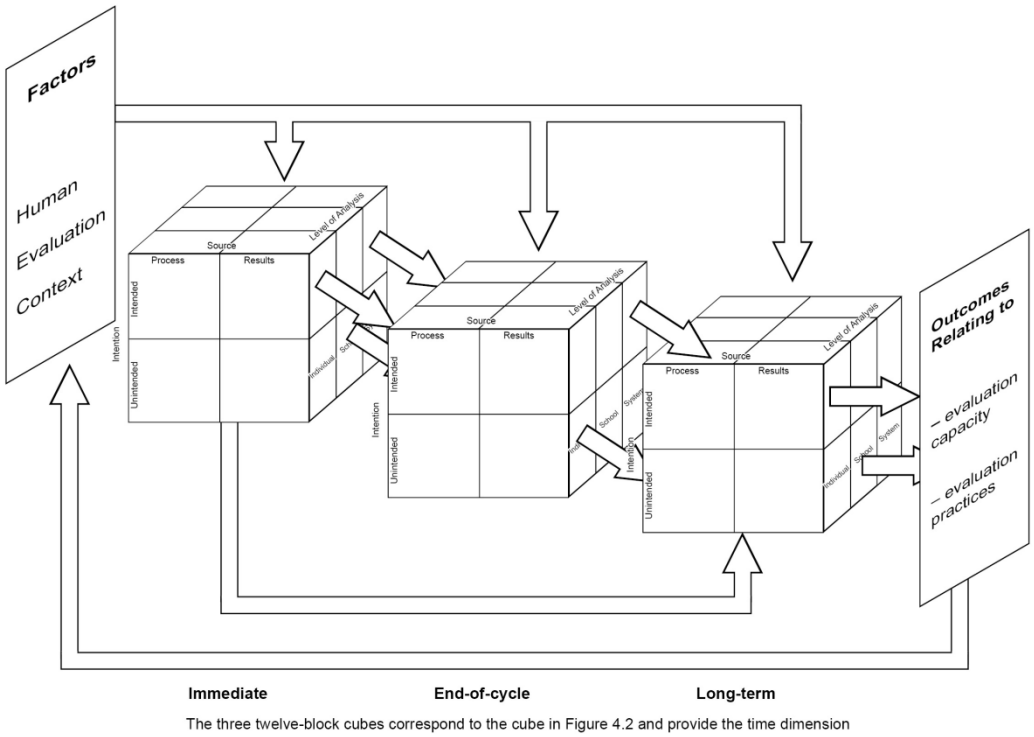
The provisional theoretical model proposed by the researcher for measuring evaluation influence in Cyclical Reviews allows any or all of the four kinds of mechanisms, described by Mark and Henry (as shown in the previous chapter in Figure 3.2)—general influence processes, cognitive and affective (or attitudinal) processes, motivational processes, and behavioural processes—to be mapped against three dimensions: the first, intention; the second, source, as described by Kirkhart (as shown in Figure 3.1); and the third, level of analysis—individual (the principal as individuals), school (the principals as leaders in their school), and system (the principals as leaders across a group of schools, at the regional level, a system level, or all three)—as described by the researcher. These three dimensions are shown in Figure 4.2.

Figure 4.2 Model showing the three dimensions of source, intention, and level of analysis (Adapted from Kirkhart, 2000:8; and Mark & Henry, 2004:41)



The model adds a fourth dimension, time, as described by the researcher (Figure 4.1). The theoretical model allows for both long-lasting and delayed influence (Kirkhart, 2000:17) as well as allowing for factors triggering influence—human (prior experiences and attributes); evaluation context, practices, and processes; and setting and environment—that emanate from the participatory nature of Cyclical Reviews to be mapped. Finally, the model shows the desired outcomes as a result of principal participation in Cyclical Reviews: improved evaluation capacity and practices in schools. The full model is shown in Figure 4.3.

Figure 4.3 Provisional Theoretical Model for the Analysis of Evaluation Influence (Adapted from Kirkhart, 2000:8; and Mark & Henry, 2004:41)



Having determined the key questions for research in Chapter 3 and proposed in this chapter a provisional theoretical model to map and analyse evaluation influence, the next chapter will explain the research procedures chosen and the rationale for their choice.

CHAPTER 5
METHODOLOGY

INTRODUCTION

The purpose of this study is to determine what influence the Cyclical Review process has on participating principals. The purpose of this chapter is to explain and justify the selection of a qualitative approach (Creswell, 2009:61–66) that uses a case-study technique based on data collected from multiple sources. The chapter therefore presents the types and sources of data collected and how their trustworthiness contributes to answering the research questions. It outlines the criteria by which the participating principals were chosen, and describes the data-recording, coding, and sorting procedure used in the study. Finally, the chapter describes how responses were mapped against the provisional theoretical model of evaluation influence developed in Chapter 4.

The issues arising from the context of the study and the literature gave rise to four research questions:

1. Is there a theoretical model that can be designed to map evaluation influence?
2. What factors, prior experiences, and understandings contribute to the influence that the involvement in Cyclical Reviews in Western Sydney Region has had on the participating principals?
3. How does participation in Cyclical Reviews in Western Sydney Region influence participating principals?
4. To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?

THE INTERPRETIVIST–CONSTRUCTIVIST PARADIGM

To obtain answers to these questions it was necessary to gather data on the Cyclical Review process and the perceptions and actions of the principals in the pilot program at various stages before, during, and after the review, and, in particular, their views on the influence of Cyclical Reviews. As the 'focus for meaning [was therefore] relative to the participants' (Cooksey & McDonald, 2011:190) an interpretivist–constructivist paradigm in the wider participatory and parallel action-research environment of the Cyclical Review pilot was deemed appropriate.

It is important to inform the reader of the approach adopted in the research because that approach reflected the researcher's explicit role as both researcher and as participant observer. Such an approach is consistent with the Participatory Action Research (PAR) model advocated by (Kemmis & McTaggart, 2008). Three attributes of PAR have been used to distinguish PAR from conventional research, 'shared ownership of research projects, community-based analysis of social problems, and an orientation toward community action' (Kemmis & McTaggart, 2008:273). These attributes were characteristic of the PAR process in which the informants were involved to develop Cyclical Reviews. The informants created the data that are examined in this study. This is illustrated in Figure 5.1 on page 122.

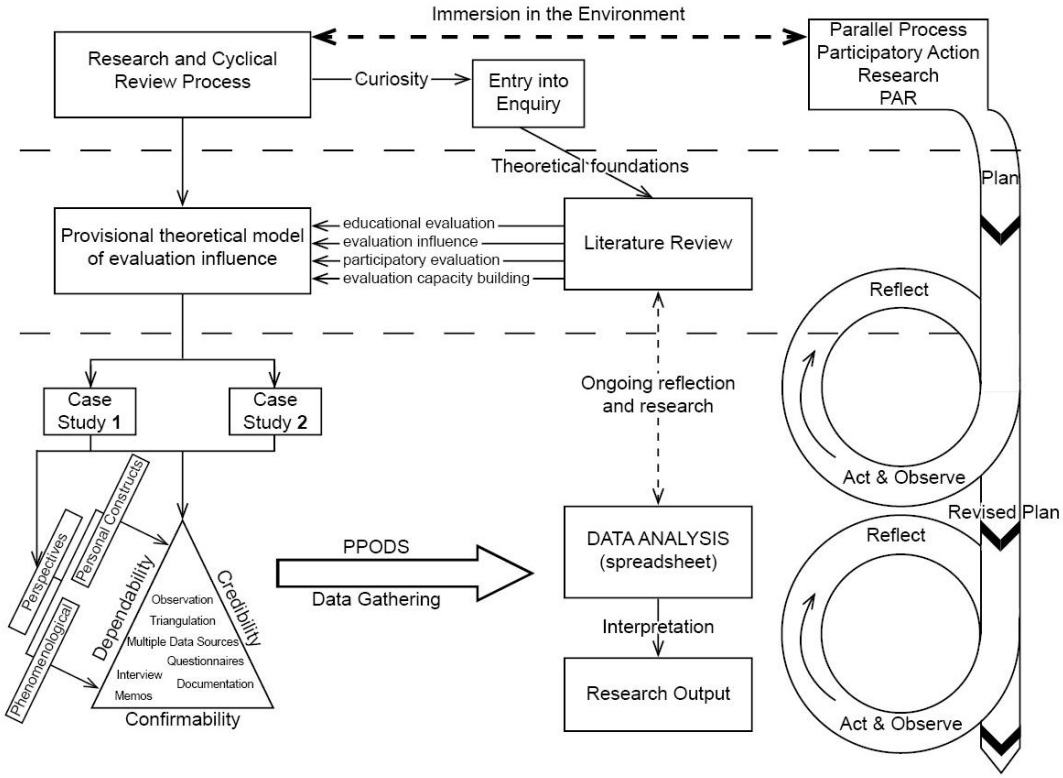
The data required to answer the questions for the study were collected using the same process (Pocket PCs for Data Collecting and Sorting, PPODS) that the principals themselves were simultaneously using to gather data about schools they were reviewing. However, these review data are not presented for the study, although they are used by the researcher in her reflection on data collection and its influence on the principals.

Given the nature of the questions (in particular, but not restricted to, the third research question, which asks 'how') and the data required to answer them, coupled with the focus on 'contemporary phenomena within a real-life context' (Yin, 2003:1), a case-study technique using multiple data sources was considered most appropriate for this paradigm and best able to align with the underpinning epistemology and ontology demonstrated in the literature review and context chapters.

Consistent with this research methodology, data were gathered using a number of qualitative methods, including participant observation, questionnaires, document analysis, and semi-structured in-depth interviews. It must be stressed that the collection of data was only possible in a context of the Cyclical Review pilot, which in turn meant that only opportunistic samples could be used.

Figure 5.1 below illustrates the research design of this study. It positions the researcher and this study parallel to and concurrent with the PAR process being undertaken to develop and implement Cyclical Reviews, as shown in the top section. Next, it shows how this study took shape from initial curiosity to exploration of the literature and the development of a provisional theoretical model, as shown in the middle section. In the bottom section, the figure shows how the provisional theoretical model was tested using case-study, outlines the sources of data that informed each part of the design, underpinned by phenomenological and personal construct perspectives, and shows the intended method of data collection and analysis based on Miles and Huberman (1994:92) and McClenaghan (2006:133), as illustrated in Figure 5.2 on page 152.

Figure 5.1 Interpretivist–Constructivist Methodological Structure



A QUALITATIVE APPROACH TO RESEARCH

At the end of the previous chapter a provisional theoretical model was developed to guide data collection and analysis. In this study the provisional theoretical model of evaluation influence served as the basis for organising the collection and analysis of data.

In choosing this approach the researcher was aware of and influenced by Smith's (1982) assertion that researchers and educationists have found the traditional quantitative research tools—for example, tests, statistics, and effect size—inadequate to decipher the interactional, political, contextual, and cultural issues in schools and

therefore inadequate to determine the influence that a review might have on a principal in the broadest sense. Denzin and Lincoln's (2008a:4) generic definition of qualitative research indicates that it is

> a situated activity that locates the observer in the world. It consists of a set of interpretive, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs, recordings, and memos to the self. At this level, qualitative research involves an interpretive, naturalistic approach to the world. This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret phenomena in terms of the meanings people bring to them.

A Phenomenological Perspective

Within a qualitative approach, people's perceptions of phenomena are important. A phenomenological perspective, according to Wiersma and Jurs (2005:243–244), has the following implications for the conduct of research.

1. As much as possible, a priori assumptions about the phenomenon under study are avoided.
2. Reality is viewed holistically and complex phenomena are not related to a few variables.
3. Data collection procedures and instruments, although having some structure, should have minimum influence on the phenomena under study.
4. There is openness to alternative explanations of the phenomenon, which may lead to alternative and changing concepts of reality.
5. Theory, as applicable, should emerge from the data as grounded theory rather than preconceived theories.

Essentially, the phenomenological perspective emphasises that reality consists of the meaning of experiences by those being studied (Wiersma & Jurs, 2005:243); that is, the meaning of reality is that which the perceiver perceives it to be. Thus, when researchers adopt a phenomenological perspective they set aside their own experiences in order to understand those of the participants in the study (Creswell, 2009).

The importance of phenomenology in modern thought was crystallised by Rogers (1964:122), a pioneer in this field, who asserted that man

> does not simply have the characteristics of a machine, he is not simply in the grip of unconscious motives; he is a person in the process of creating himself, a person who creates meaning to life, a person who embodies a dimension of subjective freedom.

The immediate relevance of Rogers's views to the present study is their insistence that the individual person can be best understood if the observer can comprehend how the world appears to the particular person whose behaviour is influenced not only by present factors but also by past experiences, memories, and perceptions of self and competence. Consequently it became apparent to this researcher that the professional and experiential backgrounds of the participants in the study should also be explored rather than limiting her understanding only to their current context and behaviours.

Adoption of this perspective made it possible to identify the different perceptions of reviews held by those principals who both led a review and whose schools were reviewed, compared with those principals who took part only in the review of another school as the principal team-member. This approach allowed the researcher to identify the essence of their experiences, as described and understood by them as participants in the study (Creswell, 2009:231). It also necessitated that she consider her own presence within the research. The intention was to expose the personal constructs of the participants.

Perspectives from Personal Construct Theory

Personal construct theory, proposed by Kelly (1955) and expanded by Bannister and Mair (1968), Bannister and Fransella (1971), Fransella (1975), Bannister and Fransella (1986), and Dalton and Dunnett (1992), added understanding of how the individual

conceptualises his or her relationship with the world of experience. Such theory suggests an explanation of how one structures and articulates experience, assuming one is actively seeking to make sense of or extend one's experience (Pope, 1980) and how one's processes are psychologically channelled by the ways in which one anticipates events. As Kelly (1970:42) puts it

> a construct system is unique to the person. It has been described as a repository of what he has learned, a statement of his intents, the values whereby he lives, the banner under which he fights . . . a theory being put to perpetual test.

Personal constructs may change over a period of time depending on the way a person successively construes events. Therefore it is useful that people reflect upon past events to understand better their current situation in the light of past experiences. Each person is a total entity in which thinking, feeling, perceiving, reacting, and responding are interactive processes (Boston, 1999). Perception is the actual attribution of personal meaning to what is experienced in the environment, involving the senses and such intellectual operations as reasoning, remembering, discriminating, and making judgements. It is essentially a process of self-perception.

Self-perception, Sergiovanni (1977) asserted, is influenced by success or failure in satisfying higher-order needs. Thus to Rogers (1964), 'self' is added to a consciousness of the phenomenological field. Where discordance occurs between the person's concepts of self, defined in this way, and emerging experience in the phenomenological field, anxiety and defensive behaviour may occur.

From a constructivist's point of view, realities are specific and constructed and depend on the individuals or groups holding them, because they are based on experiences in the local social setting (Guba & Lincoln, 1994:109–111). New experiences, however, will not result in changes in the 'self' structure if they are perceived to produce unrelieved anxiety, constrained responses, self-denial, or conclusions that are incompatible with currently held convictions about reality.

The implication of personal construct theory to this study is that the responses of the participants to questionnaires, interviews, informal discussions, the presence of the researcher, or the conduct of the study could not automatically be taken at face value. Rather, they needed to be referenced against what might have been the participants' inner motivations, their self-images, and their perceptions of the actions of those around them. This coheres with Ginsberg and Rhett's (2003) assertions that the potential users of evaluations have their own accumulation of knowledge and beliefs that need to be considered. Within her framework Kirkhart (2000) emphasised the importance of working within the social construction of user knowledge. Thus, as Leviton (2003:526) suggests:

> the context becomes central, because even the best 'body of evidence' seldom dovetails rationally with the users' knowledge. Individually, users are not *tabula rasa* [sic] and their processing of information will include many considerations over and above a single study.

THE CASE-STUDY TECHNIQUE

When a single group is to be studied in depth, qualitative research often considers using case-study (Wiersma & Jurs, 2005:17). Case-study is also used when the researcher needs to describe an entity over a period of time rather than test a hypothesis about that entity, when there is a desire to understand complex social phenomena, and when holistic and meaningful characteristics or real-life events need to be retained (Yin, 2003). That this study involved a group of principals interacting over a period of time and able to be studied in depth provided a powerful argument that case-study was an appropriate research technique.

Case-study usually follows specific stages. The first stage is determining the present complexion of the entity being studied. Background information is then sought: prior information that may be derived from literature, documentary evidence, or from

interviews with those who have been involved with the entity being researched (Denzin, 1978).

After background information has been gathered, the next stages in case-study technique are the formulation of the research problem, followed by the collection of information about the research problem and associated questions. This information may arise from many sources—official documents, records and reports, minutes of meetings, personal observation, questionnaires, and specific interviews. In this study all of these sources of information were used as the action unfolded and they in turn led to further exploration of related literature and a further refinement of the research questions, as described in Chapter 3.

As a result of the above stages, the researcher draws conclusions about the nature, activities, and performance of the entity being studied and makes recommendations about the future performances of the entity. The researcher may also generalise the findings to similar entities. In this study the entity being researched was the influence of Cyclical Reviews on the participating principals. These principals were divided into two groups, with each group being the subject of a case-study. The modest size of each group imposed severe limitations on generalisation.

> Another characteristic of the case-study approach seems to be the latitude and apparent general freedom which the investigator has with respect to the type and amount of data gathered, the source of information and procedures used to gather the information.
>
> (Helmstadler, 1970:50)

There were three reasons why this latitude offered a powerful argument for using the case-study technique in this study. First, as each pilot review was conducted, the data gathered were used to modify and improve the process. Therefore the case-study approach did not interfere with continuous improvement of the review process. Second, the complex mix of opinions, attitudes, and skills in evaluation of the

participating principals could be catered for by this technique. Third, the continuous role of the Steering Committee in suggesting and authorising the data collection could be accommodated. Flexibility in the research procedure was therefore most desirable, given this complex mix of circumstances. This technique is thus acceptable, provided the trustworthiness of the data is demonstrated.

The main advantage of case-study in comparison with other research techniques is that it enables an intensive examination of an organisation, in this case the two groups of principals, and the individuals within it. As noted by Punch:

> only an in-depth case-study can provide understanding of the important aspects of a new or persistently problematic area. This is particularly true when complex social behaviour is involved, as in much education research.
>
> (Punch, 2009:123)

A more meaningful understanding of the characteristics of the organisation can be unearthed using this technique. When a comparison is made with laboratory research, a case-study enables a more interpretative analysis of organisational phenomena (Brandt, 1982; Creswell, 2009; Diesing, 1972). Case-study provides observation in naturalistic settings, potentially unhampered by theories or predetermined conclusions (Punch, 2009; Yin, 2003). Hence the approach selected is paradigmatically interpretative in nature (Creswell, 2009; Morgan, 1980) and provides a method of inductively developing ideas from grounded data (Charmanz, 2006; Glaser & Strauss, 1967) about organisational phenomena.

Conversely, the case-study may be subject to a considerable degree of bias if the organisation being studied is one to which the researcher or the subjects of the research have attachments, obligations, and responsibilities. Unless the danger of such bias is recognised, and due allowances are made (as discussed in the section 'A Phenomenological Perspective' above), the case-study may give rise to conclusions

preconceived or intended by the researcher or the subjects of the research (Helmstadler, 1970; Stern, 1979).

To a large extent case-study relies on the impressions people have gained of an innovation or process and their recollection of their actions and the activities related to that innovation or process. In this study these impressions were gained by personal observations, questionnaires, and in-depth interviews of a selected number of principals. Case-study gives rise to the interpretation of the factors or motivations prompting the innovation and of the responses of those participants affected by it. Memories can be unreliable and personal prejudices can colour a participant's retrospective interpretation of events. Thus much of the researcher's interpretation of such data may be undermined by unreliable and subjective hindsight. Collecting the data as close as possible in time to the actual event and triangulating with data from other sources were used to mitigate against this.

A common disadvantage related to case-studies is the limited potential for generalisation, although findings can be 'generalizable to theoretical propositions and not to populations or universes' (Yin, 2003:10), as in this study. Another disadvantage can be the lack of rigour, although in this study the use of multiple sources of data was intended to control bias and provide many sources of evidence. A third common disadvantage is that case-study can result in very lengthy narratives that become unreadable (Feagin, Orum, & Sjoberg, 1991). In this study the case-studies were strictly delimited by the provisional theoretical model described in Chapter 4, Figure 4.3.

DATA SOURCES

The data required for this research were indicated by the provisional theoretical model of evaluation influence that occurs through participation in Cyclical Reviews. The scope of the data required had to include:

- factors triggering influence (human, evaluation, and context)
- the source, intention, level of analysis, and timing of the influence
- outcomes relating to evaluation practices and capacity building.

To this end a range of data-collection methods were employed. The first was documentary analysis to understand patterns of influence perceived by principals and to provide insights into factors triggering influence. Included here were documents related to the preparation for and implementation of the Cyclical Review process and documents related to specific reviews. The second was questionnaires timed to capture the principals' perceptions of immediate, end-of-cycle, and long-term influences at the time they occurred. The third was in-depth semi-structured interviews to probe and explore in greater detail the perceptions of a few principals from each case-study. The fourth was participant observation to reveal and understand emergent patterns of influence as evident in people's actions.

All of these methods used the provisional theoretical model to structure the data collected using PPODS. Before any data were collected, participating principals reached an agreement through consultation that the pilot Cyclical Review process would be evaluated and those taking part in it would contribute to the data collection at various stages over the period 2006–09 (New South Wales Department of Education and Training Western Sydney Region, June 2006). Informed consent was obtained from all participating principals at this point using the appropriate consent form required by the UNE's Human Research Ethics Committee.

Training in the use of PPODS was later provided for all participants in the Cyclical Review process. PPODS is simply an electronic process for collecting, organising and displaying data in a spreadsheet for further analysis. PPODS had been developed by the researcher for the Cyclical Review process and its use was extended for this study. This is explained in detail in the section, 'Data Collection, Coding, Sorting, and Analysis Using PPODS', later in this chapter.

All data were subjected to theme searches to help answer the research questions in terms of the provisional theoretical model developed in Chapter 4. PPODS also enabled an intensive examination of the perceptions of the participating principals. The data derived from document analysis, participant observation, questionnaires, and interviewing were entered into PPODS and this enabled their trustworthiness to be established by triangulation. These procedures will now be considered in turn, followed by a description of how the data were entered, coded, and sorted using PPODS.

Document Analysis

As asserted by Silverman (1993:89) and Punch (2009:200), documents can provide a rich and available source of data for analysis and can be examined from four perspectives: the social production of the document; the social organisation of the document; a direct analysis of the text for meaning; and the application of a theoretical perspective to the analysis of the documents.

In this study two categories of documents were examined from each of these four perspectives to establish patterns of influence. The first category related to the preparation of the Cyclical Review process and its implementation in general; and the second category related to the specific reviews in the schools of the participating principals.

Category 1: Documents related to the preparation for and implementation of the Cyclical Review process

Meeting minutes of the Working Party and Steering Committee were the two sources of documents analysed in this category. The composition of these groups, frequency

and purpose of their meetings, and the documents produced, have already been discussed in Chapter 2.

The documents were scrutinised to provide evidence about the origins and intensity of patterns of influence. The analysis also focused on both the literal meaning of the words in the minutes as well as the deeper and more symbolic meanings (Finnegan, 1996:149). Such interpretation was achievable because of the researcher's participant–observer role at the meetings.

The documentary analysis also uncovered evidence from the participating principals who were members of the original Working Party and later the Steering Committee. The influence on these principals was made explicit in their contributions to ongoing meetings and to the development of the Cyclical Review process as a form of PAR.

The minutes recorded the business of the meetings and the decisions made, and while these were adopted by the committees at the next meeting as a true and accurate record of the meeting, they were not intended always to capture every argument or opinion put forward by an individual committee member. In particular, the researcher was cognisant of the 'social organisation of the documents' in this respect, considering the issues of purpose, audience, and omissions (Hammersley & Atkinson, 1995:173). Where omissions were felt to be significant, the researcher recorded these separately as part of her ongoing participant–observer role. She then verified her observations with the committee member in question or one of the school development officers on the committee. This member-checking helped ensure the robustness and fairness of the data.

Category 2: Documents related to specific Cyclical Reviews

For the second category of documents, each of the participating principals from the first group (that is, those principals who led and hosted reviews) provided the following documents of relevance: the review report; the school plan; and annual school reports. Principals from the second group (that is, those principals who participated in reviews as team members) provided the school plan and annual school reports. All principals from the first group and six from the second also provided executive or staff-meeting minutes that directly related to the review. The absence of any reference to the review by three participating principals was also significant. These documents were made available to the researcher, although the range and depth varied from school to school. Nearly all the principals also sent emails to the researcher in her regional role about the reviews, and three sent emails that specifically related to the influence of the review. All of these forms of information were subject to documentary analysis.

The researcher did not participate in or observe the meetings that generated the school-level documents, except for the review report. As asserted by Punch (2009:201), 'documents and texts studied in isolation from their social context are deprived of their real meaning'. None the less, an understanding of these documents' social production and context was maintained through the researcher's own knowledge and familiarity of school contexts and by clarification and cross-validation in questionnaires and interviews. The documents produced by the Cyclical Review process had distinctive features.

Cyclical Review Reports. The three-day Cyclical Review culminated in the presentation of a report to the staff. The report contained achievements and future directions for the school. These reports were analysed against the intended results-based influence over time for each of the participating host-principals in the first case-study.

While each report was intended for direct use by the host school, there was also the possibility that the team-leader principal or team-member principal might have been influenced by the report he or she had helped develop, a not uncommon occurrence in PAR (Brydon-Miller, Kral, Maguire, Noffke, & Sabhlok, 2011; Macpherson, 1983). Hence these reports were also analysed for examples of unintended results-based influence over time for each of the participating lead-principals in the first case-study and the participating principals in the second case-study.

School Plans. When the pilot reviews took place, each school was expected to write a school management plan that was endorsed by the school education director. These plans needed to align with the priorities of the *State Plan*, the Department's *Strategic Plan 2006–2008*, and the *Regional Strategic Plan 2006–2008* and set targets in line with the regional targets. One outcome of the Cyclical Review process expected by the Cyclical Review Steering Committee (New South Wales Department of Education and Training Western Sydney Region, February 2008) was the modification in accordance with the review's findings of the current school plan and at least the following year's plan. At the end of 2008 the Department imposed a new planning cycle and structure. Such modifications were regarded as indicators of influence. Again, the Cyclical Review Steering Committee's expectation (New South Wales Department of Education and Training Western Sydney Region, September 2008) was that if the reviews had some medium-term to long-term influence there would be evidence of this in the new school plan. School plans and review reports of the schools where a review had been undertaken and of the schools where the principal had led the review or been a review team member were compared for such evidence.

Annual School Reports. Each government school in New South Wales is expected to publish an annual school report. The prime audience for this report is the parents. Many sections of the report are mandated, including evaluations of one area of management and one area of curriculum. The host principal was required to include a report about the Cyclical Review in the subsequent annual school report. This would

usually be included under one or both of the mandated evaluations. The direct focus on literal and deeper meaning (Finnegan, 1996) in the analysis of the evaluation section of the report pertaining to the Cyclical Review proved particularly useful in substantiating the results-based use (the upper-right-hand sectors of provisional theoretical model) of the reviews themselves.

As all principals were required to carry out curriculum and management evaluations, it was further expected, if the review process had had some influence on the evaluation capacity of principals, to find evidence of this influence in the evaluation sections of their annual school reports, regardless of the role they had played in the review. In particular, the links drawn between the methodology, the findings, and the future directions could be expected to be clearer. Bearing in mind, however, the audience for these reports, it was necessary to check how they were produced and who actually wrote this section of the report. In Punch's (2009:201) terms the 'social production' and 'social organisation' of the documents needed to be considered in order to make sense of them and limit misinterpretation.

Meeting Minutes. Executive and staff meetings are a normal feature of all the schools in the study. The meeting minutes were examined for evidence of references to Cyclical Reviews. In all schools where a review took place, the actions to be taken as a result of the review were expected to appear in both the executive and staff meeting minutes. In the team-leader and team-member principals' schools some references to Cyclical Reviews were expected to occur at least in their executive meeting minutes. It was believed the data could provide evidence for many parts of the provisional theoretical model.

A number of writers have listed a range of limitations of documentary analysis. Of relevance to this study, Creswell suggested (2009:180) that 'not all people are equally articulate and perceptive'; some documents 'may be protected from public or private access'; the 'materials may be incomplete'; and the documents 'may not be authentic or

accurate'. McCall and Simmons (1969) suggested the chief disadvantage was that the documents frequently contained only the bare essentials of events that had taken place and this deficiency could not be overcome. There is also the deceptive nature of official documents, which may be very brief or may fail to communicate the politics of the situation or the symbolism of the language used, leading to deprivation of their real meaning (Jupp, 1996).

Within the minutes of meetings, the arguments raised and minority views are not always recorded, with recognition being given only to the decisions reached by the majority (Hammersley & Atkinson, 1995). As MacDonald and Tipton (1996) indicated, what is not in the archives and not recorded in committee minutes could be just as important as what is recorded. Despite these limitations, documents are a particularly useful source of information. When they are examined from the four perspectives of social production and social organisation of the document, the direct analysis of the text for meaning, and application of different theoretical perspectives to the analysis of the document (Punch, 2009), many of these limitations can be minimised.

It is acknowledged that as the researcher was privy to and in fact contributed to the compilation of the regional-level official documentation relating to Cyclical Reviews, her commitment to such documents could be problematic to this study as a potential source of bias. To minimise this possibility, participant–observer data were usually recorded immediately following regional meetings. School-level documents also offered limited data (executive and school meeting minutes, school plans, and annual school reports) in that they did not always provide insights into the decisions and deliberations of the principals. To partially overcome this limitation, the questionnaires and the later in-depth interviews enabled expansion and clarification where the documents alone lacked detail and explanation.

Questionnaire Data from Cyclical Review Process

Multiple sources of data-collection methods are common with case-study techniques (Stake, 2006). While many case-studies use sociological and anthropological field methods such as observations and interviews, they may also use questionnaires (Punch, 2009:121). Questionnaires are often chosen as a means of obtaining information on matters not readily observable by the participant–observer, such as 'attitudes, values, opinions, and beliefs' (Cohen, Manion, & Morrison, 2007:352; Punch, 2009:248). They can have some advantages over direct observation and interviews, because they can collect survey information, provide structured (often numerical) data, are completed without the presence of the researcher, and are often relatively easy to analyse (Cohen et al., 2007:317).

This study used three questionnaires (see Appendices 1–3 pp. 296–99) to supplement participant observation for the following reasons. It was impossible to observe all actions over a two-year period of the eighteen participating principals in schools as widespread and complex as those studied. It was impossible to gauge the participating principals' motives, impressions, and opinions solely by explicit observation. It was, however, possible to collect some structured numerical data in this way. Thus while participant observation could be used at the time of the review to gauge the extent to which the participating principals understood and embraced the process, questionnaires were used to explore specifically what the principals thought about the process, what changes they would make and why, and their impressions of the influence the process and results would have on them, their school, and the region.

Where the participant–observer may be compelled to observe long periods of the same activity, the questionnaire allows the researcher to focus the collection of data on the specific areas of investigation. Questionnaires can therefore be directed to those areas of significance to the study, whereas the participant–observer using an interpretative

paradigm tends to avoid influencing the situation, but requires a much longer time to gather the required information.

To develop useful questionnaires, initial drafts of a questionnaire should be tested on individuals similar to the intended respondents (Creswell, 2009:160; Wiersma & Jurs, 2005:164). However, the degree of trialling and subsequent refinement of the questionnaire may depend on the types of questions posed. As Cohen, Manion, and Morrison (2007:321) state, 'if a closed and structured questionnaire is used, enabling patterns to be observed and comparisons to be made, then the questionnaire will need to be piloted and refined so that the final version contains as full a range of possible responses as can be reasonably foreseen'.

In the case of the three questionnaires used in this study and discussed below, and with the literature relating to questionnaire design (Cohen et al., 2007; Creswell, 2009; Oppenheim, 2001; Sudman & Bradburn, 1982) being taken onto account, all written questionnaires were subject to trials appropriate to their design and purpose, and were refined following discussions by the Working Party and later the Steering Committee. Relevance was ensured by developing an explicit rationale for each question posed. The required level of detail in responses and avoidance of confusion were ensured through trialling of and discussion about the questionnaires with principals and school development officers on the Steering Committee. Accuracy was enhanced by wording and sequencing the questions and providing a copy of the review report with the questionnaire in order to facilitate recall. Respondents were asked in successive questionnaires to reflect on influences that had occurred from previous data-collection points. This provided evidence of PAR by using 'a reciprocal process whereby each party educates the other and the intent to create local knowledge for improving the conditions and quality of life' (Miller & Crabtree, 2008:344).

Immediate questionnaire returns

From the outset, there had been an agreement between the researcher (in her regional role as a school development officer and the Executive Officer of the Steering Committee for the Cyclical Review process) and the Regional Director (Ikin, 2006) that data would be collected from a range of sources, including all participating principals, so that continuous improvements could be made to the process itself and so that information about the effects of the reviews could be collected. At this early stage, consistent with the key features of PAR (Kemmis & McTaggart, 2008:276–284), no specific instruments, questions, or hypotheses had been generated.

As suggested by Punch (2009:135–138), the Cyclical Review process was advanced using a mix of action-research methods. The focus developed as it progressed and data-collection instruments and questions were developed progressively and *in situ*. In line with this approach, the initial questionnaire (Appendix 1, p. 296) was open-ended and semi-structured, as described by Cohen, Manion and Morrison (2007:321)

> . . . between a completely open questionnaire that is akin to an open invitation to 'write what one wants' and a completely closed, completely structured questionnaire, there is the powerful tool of the semi-structured questionnaire. Here a series of questions, statements or items are presented and the respondents are asked to answer, respond to or comment on them in a way that they think best. There is a clear structure, sequence and focus, but the format is open-ended, enabling respondents to reply in their own terms. The semi-structured questionnaire sets the agenda but does not presuppose the nature of the response.

Structure, sequence, and focus were provided by asking principals to use PPODS. They could, however, record any comments they wished to make about the process in terms of the preparation, the review itself, the report, or any other area they felt was relevant. They were also asked to code these, indicating, for example, whether the comment related to the preparation for or conduct of the review and to indicate also whether they were expressing something they felt was an achievement or something they felt

needed to be addressed for themselves or in future reviews. As such, the data from these questionnaire returns provided information that directly related to the immediate influences mapped against the provisional theoretical model and were coded accordingly.

Two weeks before the formal data collection the principals were given the questionnaire and asked to trial PPODS to record their thoughts and actions about another program that was running in their schools. All reported the questionnaire was explicit and unambiguous and the device, format, and process easy to use.

End-of-cycle questionnaire returns

The questionnaire (Appendix 2, p. 297) developed for this phase was more structured than the first and was administered approximately three months after each review. This timing corresponded to the 'end-of-cycle' phase.

This questionnaire contained both closed and open-ended questions, as outlined in the research literature on questionnaire design and as discussed by Cohen et al. (2007:317–332) and Oppenheim (2001). Four of the seven closed questions asked respondents to rate their knowledge, skills, and understandings before and after the review on a six-point Likert-type scale. The remaining three questions sought the respondents' attitudes to the review: one by a rating scale and two by dichotomous responses. These questions enabled some comparisons to be made across and between the groups of principals (Oppenheim, 2001) and to provide some quick analysis so that continuous improvement could be made to the process itself. As each case was also site-specific, the remaining four questions were open-ended, seeking freer responses and enabling richer data and personalised feedback to be provided by the respondents (Cohen et al., 2007:321).

The team members from the first Cyclical Review, other than the principals themselves, completed the questionnaire approximately one month after the review. (The principals completed the questionnaire approximately three months after the review.) The team-members' responses were then examined and analysed by the Working Party and adjustments made to the questionnaire before it was given to the principals. Consistent with PAR procedures, given the small size of the sample and the qualitative richness of the data collected, the researcher decided there was no need to measure the reliability of items in these questionnaires.

Long-term questionnaire return

The third questionnaire (Appendix 3, pp. 298–99) was developed and administered in 2009 to coincide with the long-term phase of each review.

The items in the long-term questionnaire guided the respondents to reflect on the influence of the reviews from all of the four dimensions of the provisional theoretical model: influence through process and results; intended and unintended influence; levels of influence; and influence over time. As the responses were site-specific and largely unpredictable (Cohen et al., 2007:321), open-ended questions were used in this semi-structured questionnaire.

As with the end-of-cycle questionnaire, some team members, other than the principals, first trialled a draft. This draft questionnaire and the responses to the trial were discussed by the Steering Committee, with the questionnaire being modified before dissemination. The Steering Committee's action-research processes continued after this time.

There are some limitations to the use of open-ended and semi-structured questionnaires in qualitative research. The methods used by the Steering Committee

were strongly reflective of the modified-Delphi method (Wiersma & Jurs, 2005:281–291) in which an expert panel uses the progressive development of questionnaires to map understandings in a context of insufficient theory and information. In this case the effectiveness of the questionnaires was determined by the expertise of the panel in a situation where knowledge of Cyclical Reviews was emergent. Another limitation was that the panel did not systematically consider the international research literature on Cyclical Reviews or on the Delphi method. A third limitation was the turnover of about one-third of the members, and some members did not attend all the meetings. A fourth limitation was the unequal capacity of the panel members to analyse the feedback collected and explore the implications, although this was ameliorated to some extent by the researcher's coding and providing distributions of responses. There were no instances where the researcher was aware of an artificial consensus having been reached.

In summary, questionnaires were used to ascertain the actions, opinions, and impressions of all the participating principals in the study. They were used as a primary data source. As the researcher had ongoing contact with all of the principals, the questionnaires were also used as preliminary information for follow-up discussion during participant observation, or in some cases in follow-up telephone conversations or emails. Data from these episodes were recorded immediately *in situ* and analysed against the provisional theoretical model.

Interviews

Interviewing as a research method may serve a number of purposes (Punch, 2009:146). It can be the principal data-collection instrument; it can be used to test hypotheses; and it may be used in conjunction with other data-collection methods, to 'follow up unexpected results, for example, or to validate other methods, or to go deeper into the motivations of respondents and their reasons for responding as they do' (Cohen et al., 2007:351). In this study interviewing was primarily used for this third purpose. It was

designed to follow up information gained from other sources, to extend the dialogue begun with the questionnaires, and, in particular, to draw out the views about patterns of influence of those participants whose questionnaire responses raised unanticipated areas of discussion.

The type of interview that would elicit the data required was therefore an important consideration. Patton (2002) outlines three types of interview: informal conversational interviews; interview guide approaches; and standardised open-ended interviews. Fielding (1996b) describes three: standardised, semi-standardised; and non-standardised. LeCompte, Preissle and Tesch (1993) provide six categories of interviews: standardised; in-depth; ethnographic; life history; elite; and focus group. Cohen et al. (2007:351) add the non-directive interview 'in which the interviewer takes on a subordinate role. Minichiello, Aroni, Timewell, and Alexander (1990:89) and Morrison (cited in Cohen et al., 2007:358) advance continuum models. In common is the recognition of variation across two dimensions: 'the degree of structure in the interview, and how deep the interview tries to go' (Punch, 2009:145).

Connell and Campbell (2007:317) suggest that semi-structured interviews are appropriate where the research methodology needs to be flexible and adjustable during the study, as it does for case-studies. Such interviews were therefore determined as appropriate for this study. In gaining insight into individual experiences and beliefs, the researcher must be able to 'follow up on interesting leads thrown out by the interviewee'. Semi-structured interviews value the interviewee's 'personal interpretations of their experience' allowing the gathering of rich data.

It was important to look for regularities in the responses so that, if found, some generalisations from the data could be made. Equally importantly, it was also necessary to attempt to portray and capture the uniqueness, quality, and complexity of the responses and to understand why the respondents said what they said. To do this adequately a semi-structured interview was designed. In its development the

researcher was influenced by Kvale's (1996:30) description of the key characteristics of a qualitative interview. In particular, the following characteristics were taken into account:

- engaging, understanding, and interpreting the world of the respondent
- using natural language to gather and understand the respondent's knowledge
- focusing on specific ideas and themes by having direction, while not being too tightly structured
- adopting a deliberate openness to new data rather than being too prescriptive
- being able to reveal and explore the nuanced descriptions provided by the respondent.

During the course of the interviews of the four key informants, two major strengths of this method became apparent. The informants were targeted in the sense of being closely focused on substantive issues about sources and intensity of influence and they drew considerable insight about perceived causal inferences. On the other hand there were limitations due to the small number interviewed, the possibility that the open-ended questions might have created degrees of bias, occasions when inaccuracies might have crept in owing to poor recall, and occasions when interviewees might have made responses to meet what appeared to be the researcher's expectations. It was anticipated that some response bias might have been due to power differentials between the researcher and the respondents. The researcher used probes, requested examples, and used triangulation to investigate the possibility of response bias. No evidence of response bias was found.

The interviews therefore sought some preliminary information to set the respondents at ease and to confirm basic facts about them and their participation in the reviews. Respondents were then asked to talk about their involvement in and reactions to the review. Individual exploratory and probing questions were asked as the interview progressed. In addition, the researcher had prepared a set of specific questions

(Appendix 4, pp. 300–02) to be used as required during the interview and referring to each of the key themes that had emerged by this stage of the study.

Participant Observation

Participant observation is a long-established method and has been much used when researchers have been interested in the internal operations of organisations (Wolcott, 1988). According to Punch (2009:127) it gains access to the shared cultural meanings of the group members in their natural setting as they unfold. Creswell (2009:13) suggests observational and interview data are the primary sources of data in such qualitative studies. Pelto and Pelto (1976) remind us that document analysis, interview, and questionnaires are considered part of a participant–observer's approach to research. As an unstructured research method, participant observation can open up a wide field of discovery. It allows the researcher to analyse information and data as the study is proceeding, rather than postpone the analysis until all data are collected (Denzin, 1978).

Participant observation can vary in nature from complete participant to complete observer (Adler & Adler, 1994; Cooksey & McDonald, 2011; Gold, 1958; Wolcott, 1988). The complete participant acts as a member of the group under investigation but does not reveal or make overt concessions to his or her role as a researcher to any of the group members. At the other end of the continuum, the complete observer is socially distant from the group under study.

It was not possible for the researcher during the course of the study to fit exclusively into any one category of participant–observer as described by Gold (1958), Adler and Adler (1994) or Wolcott (1988). Rather, as described by Punch (2009:157), in practice 'the actual fieldwork role in research may be a blending of these possibilities'. Such blending was inevitable during the fieldwork of this study.

The factors necessitating such blending included the number of principals involved, the researcher's other regional responsibilities, the range and number of the activities, and the complexity of the issues. There was also the desire on the part of the researcher not to be seen as being too intrusive in the affairs of the schools. To attend all school-level meetings where the review might be discussed over a two-year period would have been unnecessary, too conspicuous, and logistically impossible. None the less, the researcher attended the schools of the nine principals where the reviews took place for a total of forty-five days, made contact with each of the eighteen principals a further three times, and met the principals in the study individually and in meetings a further thirty times over the course of the study. Thus an estimated eighty-three days were spent in participant observation.

The researcher in this study therefore played a blended role, which vacillated in Gold's (1958) terms between mainly participant and mainly observer, although most of the time the role carried out was at the mainly participant end of the continuum, depending on the prevailing circumstances.

The researcher was both a participant, owing to her official position (initially as school development officer and later as school education director) in the Western Sydney Region, and an observer, owing to her role as the researcher in the study. However, although an active participant in the process in her own right, she never assumed the actual positions of those under study, the principals. In this way she avoided the pretence that needs to be maintained by the complete participant, who has entered the organisation and assumed a particular role within it (Adler & Adler, 1994; Gold, 1958). Her primary role as Executive Officer of the Cyclical Review process enabled her to sustain a subsidiary role as researcher. Nevertheless, by virtue of her official position in the region, she was already part of the natural setting (Fielding, 1996a:157) and this allowed her an insider's perspective without raising some of the issues referred to by Punch (2009:158).

> They include the ethical issues associated with this method of data collection, the conceptual issues of the importance of the researcher's prior picture and the role of exploration and inspection in participant observation (Blumer, 1969), and the more practical issues of gaining access to the situation, overt versus covert observation, 'front management' or the researcher's presentation of self, and how to record what is being observed (Fielding, 1996b).

As a participant, the researcher played a leading role in coordinating and developing the regional Cyclical Review process, acted as coach to the principal team-leader on each of the pilot reviews, chaired the Working Party meetings, and initially acted as executive officer and later chaired the Steering Committee meetings. As an observer, the researcher was able to gather first-hand information from the participating principals because of her legitimate participation in the preparation for and conduct of each of the reviews.

Continual involvement with some of the participating principals through the Working Party and Steering Committee provided opportunity for participant-as-observer and observer-as-participant roles (Gold, 1958). Through these meetings the researcher had an ongoing opportunity to participate fully in the discussion, to observe the principals, and to hear their views on the evolving process and its ongoing influence. This was particularly useful in gathering data that related to influences at the collective level of analysis, as described by Henry and Mark (2003; Mark & Henry, 2004).

The closer to complete participation the observer is able to be, the more likely he or she is to become familiar with the full range of functions and influences in the organisation (Spradley, 1980) and to be given access to documents, information, and confidences relating to the organisation. In this study, the key issue was to understand what it was like to be a participating principal in a Cyclical Review. By maintaining a mainly participant focus, the researcher was able to gain initial and potentially valuable insights into how it felt to be part of the process and was able to share with the participating principals the influences the reviews had on them (Schatzman & Strauss, 1973:62). This was particularly useful at the personal level of analysis (Henry & Mark,

2003; Mark & Henry, 2004), as the researcher gained knowledge of and developed attitudes about the review process that corresponded to the team-leader principals' knowledge and attitudes.

The technique of participant observation enjoys several advantages (De Walt & DeWalt, 2002; Hammersley & Atkinson, 1995; Spradley, 1980), with the most important being that the participant–observer is usually not totally bound in fieldwork by any prejudgements about the nature of the problem. Certainly, as Angus (1987) has indicated, there was a need for a foreshadowed problem, which provided a focus for the research as a whole. Further, the problem needed to be capable of modification and redefinition as the research progressed. As it transpired, the issues, and questions that arose from them, evolved during the researcher's association with the principals concerned. To assist modification and redefinition, a less-directive approach to interviewing was adopted, whereby questions were not decided beforehand, although the researcher had a list of issues to be covered (Hammersley & Atkinson, 1995). As explained by DeWalt and DeWalt (2002:8), participant observation encouraged the 'formulation of new research questions and hypotheses grounded in on-the-scene observation'.

Conversely, it can be argued that participant observation can lead to the observer's becoming over-familiar with those under observation (Pelto & Pelto, 1976; Wolcott, 1988). Objectivity can be lost or tempered by an unwillingness to offend those with whom one has become personally acquainted. Furthermore, the act of participation can lead to an unconscious attachment to the values and procedures promoted by those under observation. Hence prejudgements about the problem avoided by the adoption of participant observation may be replaced by biased judgements arising from personal involvement. The key point is that the researcher should be aware of the compromises in access, objectivity, and community expectations made at any particular place on the participant–observer continuum (De Walt & DeWalt, 2002:23).

Again, with this study both prejudgements and biased judgements needed to be vigilantly guarded against because of the researcher's professional responsibilities. For instance, it could have been presumed that all the participating principals in the group that both led and hosted a review were equally well-versed in and committed to the content of the *Exemplary Practice Statements*, which underpinned the reviews. Equally, it could have been preconceived that the team-member principals would be far less committed to these statements at least in the early stages of the review process. Neither of these judgements would have been entirely correct.

The researcher also experienced the feeling of obligation that arises from participation both to the group and to the process. This was manifested in an urge to contribute, cooperate, and attend to routine as well as professional tasks associated with specific reviews that were beyond her official role. As a number of writers in the field (Adler & Adler, 1994; Punch, 2009; Rist, 1977) have put it, the temptation was to 'go native' and to assist with the process of review implementation in specific schools. This temptation had to be resisted. Ideally, the participant–observer would enjoy having no prejudgements about the nature of the problem. This cannot be claimed as an advantage in this study. Because of her responsibilities in the development and promulgation of the process itself, the researcher had been, in a sense, a participant–observer from the very beginning when the process was conceived to the extent that both the study and the process became inextricably linked.

Such participation and observation inevitably influenced the perceptions the researcher brought to the study, particularly as she met with the principals in the ongoing Steering Committee meetings, although it was impossible for her to manipulate events and introduce observer bias (Yin, 2003:86). With the intertwined nature of the process and the study it was impossible to divorce entirely one from the other; likewise the role of researcher from that of a senior regional officer with a responsibility for the Cyclical Review process.

It can be said, however, that many matters arising in the course of the study were not preconceived by the researcher. In Angus's (1987) terms, the researcher was confronted with changes to the predicted problem. The principals themselves disclosed many considerations at the personal, interpersonal, and collective levels that she could never have anticipated from her regional position. As a result, the initial problem evolved in scope, proportion, and complexity. The researcher thus believes the level of data trustworthiness increased as the study progressed and can be given a reasonably high degree of trust.

Data Collection, Coding, Sorting, and Analysis Using PPODS

Data collection and analysis in qualitative research are characterised by diversity (Coffey & Atkinson, 1996:14; Denzin & Lincoln, 2008b:4; Miles & Huberman, 1994:9; Punch, 2009:170; Ryan & Bernard, 2003a; Tesch, 1990:3–4). Punch (2009:171) asserts, 'there is no single right way to do qualitative data analysis—no single methodological framework'. Coffey and Atkinson (1996:3) stress that 'what links all the approaches is a central concern with transforming and interpreting qualitative data—in a rigorous and scholarly way—in order to capture the complexities of the social worlds we seek to explain'. 'If the researcher needs to invent, or piece together, new tools or techniques, he or she will do so'(Denzin & Lincoln, 2008b:5).

Despite the variety of approaches, some writers have sought to identify common features of or practices in qualitative data analysis. Tesch (1990:95–97) identified ten principles that she believed held true for most types of qualitative analysis. They are:

- data collection and analysis are concurrent or cyclic
- the analysis process is systematic and comprehensive but not rigid or inflexible
- attending to data includes a reflective process resulting in notes or 'memos' (mainly conceptual)
- data are segmented into categories

- data segments are categorised and coded according to an organising system predominantly derived from the data themselves
- the main intellectual tool is comparison (constant comparative method)
- categories for sorting are tentative and preliminary in the beginning and they continue to remain flexible
- manipulating qualitative data during coding, organising, and analysing is an eclectic activity
- these procedures are neither scientific nor mechanistic
- the result of the analysis is some type of higher-level synthesis.

(adapted from Tesch, 1990)

Miles and Huberman (1994:8) identified six common general features for generating meaning in qualitative analysis. They are:

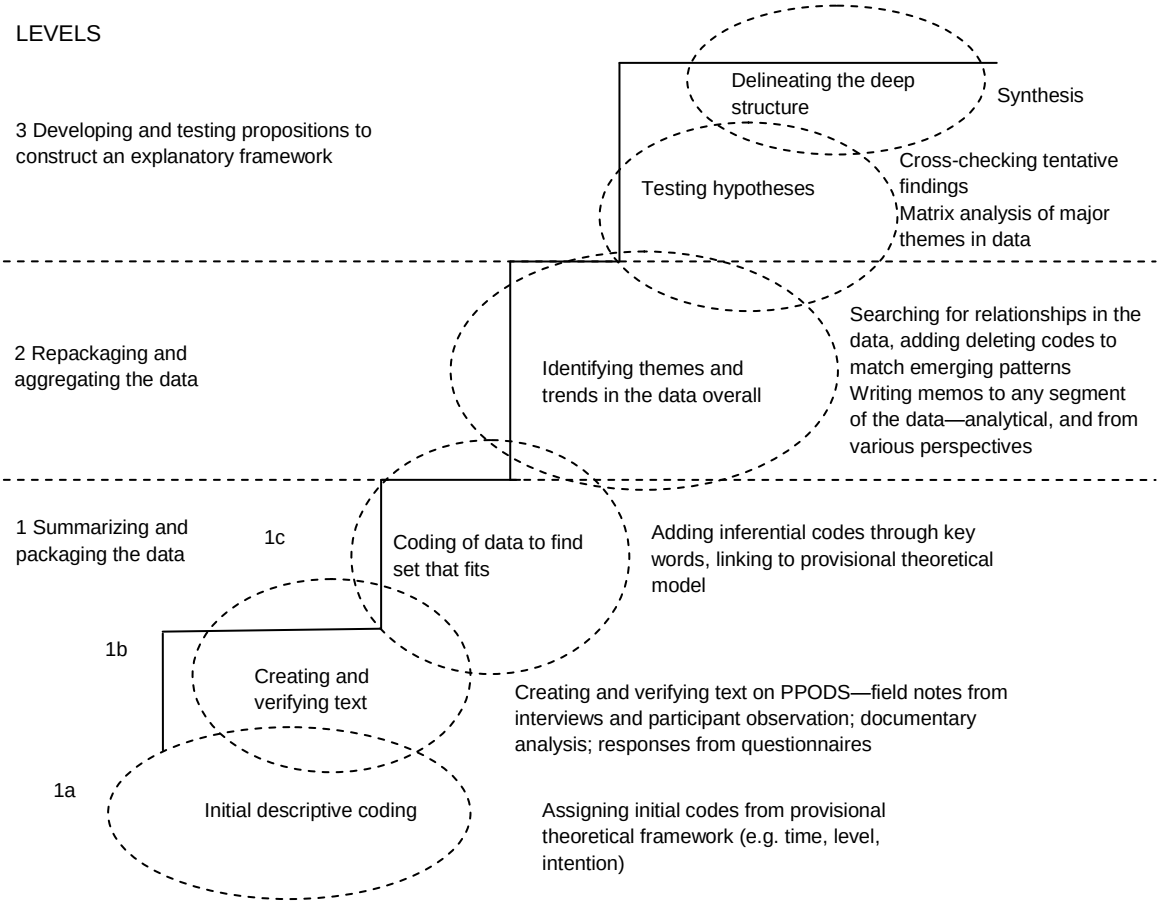
- affixing codes to field notes drawn from observations or interviews
- noting reflections or other remarks
- sorting and sifting through the materials to identify similar phrases, relationships between variables, patterns, themes, distinct differences between subgroup, and common sequences (constant comparative method)
- isolating these patterns and processes, commonalities, and differences, and taking them out into the field in the next wave of data collection (testing and validating)
- gradually elaborating a small set of generalisations that cover the consistencies discerned in the database
- confronting those generalisations with a formalised body of knowledge in the form of constructs or theories.

The researcher took Tesch's and Miles and Huberman's principles into account (allowing that in some instances the principles overlap) when designing PPODS. In addition, the design was influenced by the interactive model for the analysis of qualitative data proposed by Miles and Huberman (1994:12). In this model the three main components of data analysis—data reduction, data display, and drawing and

verifying conclusions—are seen as three concurrent activities, interacting throughout the analysis.

Figure 5.2 below illustrates this analytical strategy. It represents a traditional qualitative data analysis process and suits the interpretivist–constructivist nature of the current research.

Figure 5.2 The Ladder of Analytical Abstraction (Adapted from Miles & Huberman, 1994:92 and McClenaghan, 2006:133)



Following the stages in Figure 5.2 above, each data set (field notes from interviews and participant observation; documentary analysis; and responses from questionnaires) was collected, verified, entered, sorted, and analysed. Gaps and emphases in the data

were followed up through the iterative nature of the data collection process. Data sets were separated into segments and individually coded line by line, a time-consuming but accepted method (Creswell, 2009:188; Ryan & Bernard, 2003a; 2003b:188), using mainly repetition, similarities and differences, cutting and sorting, and keyword-in-context (KWIC) techniques to identify themes, based on these theme-identification techniques, as described by Ryan and Bernard (2003b). For example, any comments pertaining to a change in the principal's skill level in evaluation processes was inferentially coded first as 'behavioural' (keyword 1) and then as 'skill performance' (keyword 2). At the initial data collection point, descriptive codes (denoting time, type, intention, level, participant, review, and data source) were able to be applied to place the data within the provisional theoretical model. Notes could also be added (perspectives, comments) to provide further insight into perspectives that were relevant to the study. Figure 5.3 below shows one example of the data as they were displayed in the spreadsheet following PPODS.

Figure 5.3 Example of data item collected and entered into the spreadsheet following PPODS

Time	Type	Intention	Level	Keyword 2	Data	Participant	Review	Data Source	Key Word 1	Perspective	Comment
a	p	i	p	skill perfo	Leader was able to talk in depth about what the data meant and lead the discussion to analyse the data	El	F	o	Behavioural	Had said that she didn't really understand the data in the past but wanted to learn. Data analysis basic. Liked idea of evaluation. Was open to being reviewed.	Had grown up with inspections. Had some preconceived ideas of what a CR could do and did not always catch onto what was required.

While not strictly following the rules of inductive analysis (Patton, 1980), the researcher was seeking the 'patterns, themes, and categories' (Patton, 1980:306) that emerged from the data. Thus by using PPODS the researcher was able to:

- collect and enter data from each source using the same instrument and process
- enter data once when they occurred
- verify interview data with the respondent at the time of data collection

- use initial descriptive codes at the time of data collection
- add inferential codes through keywords at the beginning of a segment or through the addition of a column to label each segment
- add and delete codes as patterns emerged
- add memos to any segment of data
- sort and shift through data using codes and keywords linked to the sorting features of the Excel spreadsheet to compare, group, and combine data in multiple ways.

METHODS USED TO ASSURE TRUSTWORTHINESS OF DATA

Sampling

The primary purpose of this study was to consider the influences that Cyclical Reviews had over time on the principals who participated in them and, in so doing, describe these influences. The purpose of this section is to link the purposeful and opportunistic selection of principals to the trustworthiness of data.

In order to do this adequately using a case-study approach, three features were necessary. First, a 'panel study' (Wiersma & Jurs, 2005:161), whereby data were collected from the same principals over a two-year period, was adopted in order to investigate not only changes but also to identify the source of these changes. Second, a collective case-study, as described by Stake (1994), was used in order to focus both within and across cases, in this study across the two groups of principals. Third, to be able to make a preliminary degree of generalisation (Creswell, 2009:192–193; Punch, 2009:121–123; Yin, 2003), purposeful sampling was used to select principals who had sufficiently varied experiences in schools and in school reviews. It was also necessary to select principals from schools exhibiting a variety of characteristics.

Therefore all eighteen principals who participated in the nine Cyclical Reviews were included in the study and were divided into two groups of nine, and these groups became the basis for two case-studies. The first group comprised those principals who both led a review and had a review in their school; and the second group comprised those who took part in a review as a team member.

As described in Chapter 2, the nine schools in which the reviews took place were located in two of the eight school education groups that constitute the region and hence the principals who belonged to the first group also worked in these schools. The schools in these two school education groups had been involved in a series of reviews in 2002 and 2003, proposed by the superintendent of the time, who became the Regional Director in 2005 and subsequently proposed Cyclical Reviews for the region. The attitudes to reviews by nearly all of the principals and many of the staff in the schools in these two groups were already positive. There were exceptions, however, including two principals who were cautious of the benefits of a review, and staff in one school who were opposed to a review of their school. The principals in this first group not only hosted a review in their own schools and led a review in another of the schools, but also played a role in the development of the process itself.

The principals from the second group, however, who were chosen as team members, each came from one of the other six school education groups in the region. They did not play any prior role in the development of the process, did not host a review in their own school, and did not lead a review. None of these principals had been involved in the 2002–03 review process, although some of them had been involved in other types of reviews in the past. As a group they had no particular positive or negative attitudes towards reviews, although individually they displayed a range of attitudes.

To determine the membership of this second group of principals, the researcher, in her role as coordinator of the Cyclical Review process, initially made contact with the school education directors in charge of each of the six school education groups. She

explained the role of the principal team-member to them and sought their services in identifying principals in their school education groups who they considered could serve in this role. Nominated principals were then approached directly by the researcher to seek their participation in a review and hence in this study. Only one principal of those approached declined to take part, the reason given being that personal commitments were already taking him away from his school and he could not commit to further days away at that stage.

There were also a number of variations within the groups. In each group there were principals of large and small primary schools, principals of secondary schools; principals relatively new to the role and those with many years of experience; and male and female principals. The principals came from schools in both low-socioeconomic-status areas and from the more affluent areas of the region. It was important for two reasons to have a mix of school types and principal experience: first, as this was a regional pilot it was necessary to ensure that the Cyclical Review process was suitable for all schools in the region; and second, for the purpose of this study, a range of principals' backgrounds and experiences was necessary to allow for generalisations of any findings that might emerge. It should be noted that principals are referred to by gender (that is, he and she) in this study, but gender played no role in the analysis.

By initially selecting for the study all eighteen principals who were to take part in the pilot Cyclical Reviews it was expected that the sample would be large and diverse enough to sustain two case-studies. The sample size also allowed enough flexibility should a small number of principals drop out of the study, through retirement, promotion to another school, or for other unforeseen reasons.

Finally, four of the eighteen principals—two principals from each group—were selected for more in-depth follow-up interviews as 'key informants' (Kvale, 1996; Punch, 2009) during the final data-collection stage. As Wollcott (1988:195) explained, a key informant is 'an individual in whom one invests a disproportionate amount of

time because that individual appears to be particularly well-informed, articulate, approachable or available'. To select these principals, consideration was given to the earlier observations and data collection. The four principals whose responses were considered more illuminating and in-depth were considered. For instance, their answers to some questions suggested that further explanation by way of interview would unearth information not forthcoming from other sources. The range of characteristics brought by them to the study was then considered in determining the final four.

This detailed description of selection processes was intended to clarify the extent to which sampling added to the trustworthiness of the data. It was expected that the first case-study group with prior and positive experience and commitment would actively arbitrate the quality of data about the nature and sources of maximum influence. It was expected that the second case-study group would actively arbitrate data on influences in circumstances where influence was minimal. It was expected that the key informants would deepen the qualitative richness of the data obtained from both groups. Overall it was expected that these methods of interrogating the data would increase their credibility and reliability, and to a lesser degree their transferability, dependability, and comparability.

To be explicit, when Guba and Lincoln (1989) clarified these terms for qualitative research they noted (p. 241) that transferability was to be primarily indicated by the quality of 'thick description' provided. In Chapters 6 and 7 the words of the participating principals will be used extensively to transfer the essence of their thinking about influences and factors triggering influence. When Guba and Lincoln clarified dependability (p. 242), they referred to the stability of data and the increasingly sophisticated methodological constructions over time. The latter condition has already been evident in the progressive reconstruction of questionnaires and will be evident in the analysis of changes of perceptions of influences using the provisional theoretical model. Guba and Lincoln's criteria of confirmability (p. 242) emphasised the

importance of being able to track data, interpretations, and practical outcomes back to people in context. In this study all findings can be matched to sources and cross-referenced to the provisional theoretical model of influence. Their authenticity audit (p. 245) is concerned with fairness, defined as honouring the rights of stakeholders. The negotiated pathway to the development of the Cyclical Review process used in this study has been described in detail above. This pathway has led to the reconciliation of stakeholder interests and, within that context, the development of a multiple-method approach.

Even with allowance for the use of triangulation to control bias in data, there remained the difficulty of bias on the part of the principals during the interviews, meetings, and observations. Also possible was bias brought to the study by the researcher herself. As noted above, this latter difficulty may have been intensified by the researcher's being seen as an authority figure in the eyes of those principals under study or being seen as a doctoral researcher with special expertise or both. This combination also gave scope for bias and distortion to arise in the researchers' interpretation of the actions and responses of the participants and highlighted the need to guard against it.

To this end it is contended that a range of measures have been applied to ensure an acceptable measure of trustworthiness. First, each principal who took part in the Cyclical Reviews was subject to data-gathering for the study. While the researcher held a senior position in the region from 2006 until July 2008, she was not the supervisor of any of the principals in the study, and from July 2008 none of the team-leaders and host principals and only one of the team-member principals were under the direct supervision of the researcher. Furthermore, because they, and the schools they represented, were guaranteed anonymity through a data-collection and data-sorting process with which they were personally familiar and in which they trusted, it is believed that, in general, a frankness and openness ensued.

A separate question concerns the reliability of the researcher given her dual role of being instrumental to both the conduct of the Cyclical Review process and to the conduct of this study. In order to enhance reliability, the methodology was developed and defended before data-gathering. In the case of the documents analysed, none was compiled specifically for this study, the documents being for a much wider audience than the researcher alone. Further, the data-gathering by survey was controlled by the Steering Committee with input by the principals and researcher, this being consistent with action research to advance the cumulative trustworthiness of data and the questionnaires.

The reliance on school-level minutes of meetings rather than attendance at such meetings may have assisted with the trustworthiness of the data and the reliability of the data-collection methods. At such meetings, the existence of the study would have been far less conspicuous than if the researcher had been present. The discussions and deliberations of the principals at these meetings would have been uninfluenced by any consideration of the presence of, or contribution by, the researcher, and yet references or lack of them to Cyclical Reviews and consensus on actions to be taken after the principals' participation in the review were recorded in the minutes and many were able to be verified in the school plans and annual school reports.

Bias was also controlled by triangulation. No single data-collection method or data source can cover adequately all of the evaluation influences that interact within an organisation. Further, adoption of a single procedure or collection of data from a single source could render the analysis biased and distorted. When two or more different research strategies in the study of the same empirical units are used, the process is known as 'triangulation' (Denzin, 1978:301–308).

> Triangulation is qualitative cross-validation. It assesses the sufficiency of the data according to the convergence of multiple data sources or multiple data-collection procedures.
>
> (Wiersma & Jurs, 2005:256)

Further elements of trustworthiness were evident in the methods used to ensure the credibility and reliability of data. Thus data were collected from all participating principals on a number of occasions to demonstrate within-method verification. First, the researcher observed the principals over the three days of the review. Second, they were surveyed on three occasions by written questionnaires. Third, to allow a deeper exploration of salient issues, a smaller sample of participating principals were also interviewed. Using PPODS allowed these principals immediate access to their responses and hence the opportunities to challenge, correct, and explain their recorded responses. Fourth, against their individual interpretations of the influence of the reviews were matched official school and regional records, as revealed by document searches. The individual interpretations were then matched against the collective interpretations to find recurring themes and pathways. This mix of methods provided between-method verification.

The convergence of evidence from multiple sources and multiple forms of evidence were used to build a coherent justification for the emergent themes (Creswell, 2009:191; Creswell & Miller, 2000). These analyses were presented to ongoing Steering Committee meetings for challenge and comment. In these ways triangulation was therefore expected to minimise bias and thus increase dependability of the findings.

SUMMARY

This study set out to determine the influence that involvement in the Cyclical Reviews had on the participating school principals. It was expected that the study would gather the principals' perceptions of the factors, both personal and contextual, that shaped that influence and which in time would lead to heightened or lessened commitment to and advocacy of Cyclical Reviews across the system.

From consideration of the historical, political, and practical contexts of the study, as outlined in Chapter 2 and the literature of research related to the study, a set of key questions was formulated in Chapter 3. These questions gave direction to the qualitative approach adopted, the selection of case-study techniques, and a range of data-gathering methods chosen as appropriate for this study and described in this chapter. It was emphasised that while the Steering Committee advanced the Cyclical Review process using PAR, this study drew data from that context to conduct two case-studies.

The case-study technique within a multiple-method, qualitative approach was chosen as the best means of enabling an intensive examination of patterns of influence. The case-study technique also allowed the researcher to unravel some of the complexities of influence and to describe each of the contexts and the influences of the Cyclical Reviews on the principals.

The researcher was a participant and observer within the Cyclical Reviews process, as well as the study itself. Triangulation was chosen as a means of collecting and reconciling data from many sources (documents, questionnaires, and interviews) rather than relying alone on any one source. In this regard the possibility of bias was acknowledged and measures were taken to minimise it.

The sampling strategy was clarified to indicate the extent to which it affected the trustworthiness of data collected, especially in regards to case-studies. Other key criteria were considered such as transferability, dependability, confirmability, and authenticity.

The methodology and associated data-gathering methods described in this chapter were thus designed to elicit useful responses to the key research questions. The next two chapters present findings in line with the provisional theoretical model. Chapter 6

details the findings and implications about the factors contributing to influence and Chapter 7 details the findings and implications about the influences themselves.

CHAPTER 6
FINDINGS AND IMPLICATIONS ABOUT FACTORS CONTRIBUTING TO INFLUENCE

INTRODUCTION

This chapter presents an analysis of the findings and a synthesis of the implications of the factors that were found to contribute to the influences on the participating principals. The implications are presented in terms of theory, policy, and practice, discussed in Chapters 2 and 3.

FACTORS CONTRIBUTING TO INFLUENCE

The principals' perception was that several factors triggered their use of the results and processes of Cyclical Reviews and affected the influence that Cyclical Reviews had on them. These factors fell into three groups: human factors; evaluation factors; and context factors, which aligned with Alkin's (1985) groupings. Human factors included motivation and prior experiences, leadership, host principal's expectations and support, and human relationships. Evaluation factors included the data-collection processes (survey, interviews, desk audit, and observations), PPODS (overall data-analysis process, pocket PCs, codes, and data analysis), and the structures and resources (training, *Exemplary Practice Statements*, review report, and timetable). Finally, the context factors included school culture, timing of the review, team size and composition, and review focus. The data concerning these three groups will now be discussed in turn and these are shown diagrammatically at the end of each group.

Tentative conclusions are drawn after each discussion regarding these groupings and underlying values that serve as catalytic conditions.

Human Factors

Motivations

The first group of human factors to emerge from principals in both groups were the motivations that principals brought to the Cyclical Review process. From observations, the first questionnaire, and comments made by the principals during the development of the materials and the preparation for each review, the most commonly cited motivation by principals in both groups for participating in the Cyclical Review process was professional development. Whether participating as host, team leader, or team member, nearly all principals stated that they expected to learn from the review process and acquire new skills.

Indicative of responses from Group 1, one principal explained:

> *I was aware of the previous district reviews and was on a team. I see it as great professional learning. I am keen to be part of a similar action.*
>
> (Group 1, Principal B)

Group 2 responses were also similar, for example:

> *I have been working on school evaluations, making them more meaningful, tighter. I am looking forward to working with a team to see the mechanics of a sound evaluation process and learn from it.*
>
> (Group 2, Principal O)

Additionally, six principals from Group 1 also stated that they were looking forward to learning more about their school and finding ways to improve their students' learning. For example, a typical comment was:

> *I want open honest feedback and things we can work on to develop . . . to improve student learning.*
>
> (Group 1, Principal E)

Two principals from Group 1, however, seemed more intent on using the review to drive the changes they already saw as necessary, one explaining that he wanted to

drive change using data so that staff could not argue and the other explaining that he wanted the review to drive and justify the things he had already decided needed doing.

Giving rise to some of these motivations were the principals' prior experiences. Principals in Group 1 regularly referred to their experiences in 2003 in the district where the regional director, as the then superintendent, had implemented a review process. For example:

> *I trust the regional director and ——. —— led the reviews in 2003 and now is part of this process as a principal . . . gives me confidence.*
>
> (Group 1, Principal B)

Twelve of the principals in these two groups had previously been involved in Quality Assurance reviews and, although they voiced mixed opinions about the Quality Assurance period, referred to the assessing of a school using on an agreed set of statements and by a team as good features and ones that would motivate them to participate and take notice of the findings. There was some caution expressed by four of these principals, however, about getting into yet another new scheme that promised a lot but might well go the way of many previous schemes and simply disappear before there was time to prove whether it was successful or not, just as Quality Assurance never got beyond the first round. While most of the principals had been inspected for placement on a promotion list prior to 1990, only one principal compared this experience specifically with Cyclical Reviews:

> *I was around with inspections. The expertise with principals and school development officers is the good part [of Cyclical Reviews]. Same as if you got a good inspector. But it is also good to have teachers at the chalkface involved . . . gives some fresh ideas. That wouldn't have happened in the days of inspection.*
>
> (Group 1, Principal B)

Fifteen principals across the two groups did, however, comment from prior experience on their preference for a developmental process rather than one of accountability. For example, one stated:

> *Accountability, no way. I am keen on the developmental process. [I want to be] involved to further self-evaluation, evaluation, development, and involvement of my staff. [I] want the participatory and open practices. No excuses. No hidden agendas.*
>
> (Group 1, Principal A)

A telling comment came from one principal who had experienced both the inspection system and the Quality Assurance reviews and had been a team member on a number of reviews under the *School Development Policy*. She stated:

> *Some people would see the Cyclical Reviews from a developmental perspective. Some from an accountability one. As it is voluntary, the principals are comfortable and so see it as professional learning . . . development. The hosts also wrote the instruments and procedures and feel they own it. But from a teacher's perspective it is accountability. These perceptions are important for later use of the results in the school. For me it is a mixture.*
>
> (Group 2, Principal R)

Leadership of the review team

The second human factor to emerge was leadership of the review team. Twelve principals across the two groups commented that having a principal lead the review team was a motivating factor, as they perceived a principal to be non-threatening, but these twelve principals believed it was even more important to have the right principal or even another person, one who fitted the review, to lead. Six principals from Group 1 and two from Group 2 also commented that the coach and team leader formed a crucial leadership team. Principal A from Group 1 as a host principal summarised this attitude when he commented that:

> *The model of Cyclical Reviews is ideal, non-threatening, [but it] must have expertise. It is critical that the principals who lead have skills, are credible, have respect . . . know the*

challenges. The relationship [for the host principal] with the lead principal is most crucial. Therefore it [the leader] could be a school education director. Team leader and coach need complementary skills.

In their team-leader roles, three Group 1 principals who knew the host school and principal well or spent time learning about the school and who built a relationship with the host principal and took their role very seriously were regarded as having a positive influence on other principals' attitudes towards the review. They elicited comments such as '*[the] leader's understanding of the data and data analysis instilled confidence and acceptance*' (Group 2, Principal L), and '*the school was considerably advantaged in having —— lead the team . . . [he or she] knew [the] school*' (Group 2, Principal K).

Team-member principals were extremely supportive of colleagues who they believed had excellent leadership skills. For example, in one case the team leader (Principal E) commented that she made sure the codes were clearly defined at the in-school training session to match the context and ensure that participants had a common understanding of the task so that in her words '*everyone was on the same page*'. The team-member principal on this team (Principal O) saw this as strong leadership, commenting on the clarity that ensued by having a group discussion about the data and the meaning of the codes. On the other hand team-member principals were critical in those instances where they believed the leader did not have the requisite leadership or technical skills, suggesting that this was a hindrance. For example:

[The] leader needs to ensure comments are data driven. Sometimes it was based on what he thought. I would want to completely trust a leader's grasp of the process. I would be cautious about who was the leader of a review in my school.

(Group 2, Principal M)

Leadership by the host principal

The third human factor to emerge was the leadership exhibited by the host principal. Seven host principals stated that they expected the review to make a difference or be a success. In explaining this optimism, Principal A explained that he knew what they were getting into in terms of the process and, as he knew his school well, was confident the results would be positive. As he further explained, if the results of review had been unexpected then the response to the review might have been negative. The team-leader principal of this review confirmed the need for positive expectations of and support provided by the host.

> *The attitude of the host really helped. Very supportive. He wanted the process to work and to get honest feedback. I would need to ensure this if I had one at my school.*
>
> (Group 1, Principal B)

Where principals perceived, however, that the host principal's expectations or support were not at the level they expected, they had concerns about the review's ongoing influence. For example, from team-leader and team-member principals of the same review:

> *While the principal was optimistic about the review, his support in ensuring the review went smoothly could have been better. His preparation was a bit haphazard and so time was wasted when the timetable was not thought out well, rooms not specified, and so on. This actually detracted from the amount of work we could do, how much data we collected, how much time we had to analyse. Ultimately, I guess the quality, thoroughness, usefulness of the report. . . . It emphasised for me that the role of school and [host] principal reeds to be clearly stated.*
>
> (Group 1, Principal A)

> *[I] felt unsure about expectations. Principal [C] did not really come and talk to us about what he wanted or was feeling. This was a negative aspect of the process from my point of view. The findings could have been more explicit if we had a better idea of the principal's expectations. He was supportive but not in an active way. I don't think our leader was informed any more than we were. If he was he didn't say so.*
>
> (Group 2, Principal L)

Principal S from Group 2, who believed that the host principals' expectations and support were crucial to the ongoing influence of Cyclical Reviews in the reviewed school, stressed that host principals should be required to explain why they want a Cyclical Review and how they are going to use it at the time of requesting a review for their school.

Human relationships

The fourth factor to emerge was human relationships. Principals from both groups agreed that good team dynamics were essential, not only for the Cyclical Review process to be influential, but also for the results to be influential.

> *[The team were] genuinely keen to participate, learn about [the] school and evaluation . . . [this] helped a lot . . . especially the last day. [The team] got stuck into the . . . very rewarding seeing it come together. Team building and its importance came home to me. This will make a difference to some things I do back at my school.*
>
> (Group 1, Principal B)

> *[The] team worked very well together. You could sense it. That is really important because they are providing external eyes, independence . . . So they need to be thorough, inclusive, professional, highly credible. Staff are more willing to accept the report that way.*
>
> (Group 1, Principal C)

Thirteen principals commented on the satisfaction they gained from working on a professional team with another principal and a good balance of expertise, one adding that '*this factor alone is enough to make me want to do this again or have one in my school*' (Group 2, Principal Q).

When it was perceived that the team lacked cohesiveness, this was commented on as a hindering factor. Principal A, as a team leader, noted that some of the team members did not appear to be up to the task. One result observed by the researcher was that the

leader needed to do a lot of work between Days 2 and 3 to prepare the data, and the principal's perception was that the team was cohesive. In another review it was noted by the team leader (Principal G) and confirmed by the researcher that:

> *The review team was not all that cohesive. No real learning environment was created. One team member was often off task, another worked very intensely but as an individual. Made the report writing a bit difficult, [it was] hard work dragging out the findings.*

Six principals involved in these reviews commented that the reviews had confirmed to them that there was a need to look at training, competencies, or principals' recommendations for team members and that not everyone was suitable for a review team. Principals A and G also commented that this type of experience could deter some principals from future involvement. It also confirmed to these principals that owing to the short three-day timeframe for the review there was a need to develop a team culture that enabled team leaders to get to know each team member prior to the review's commencement.

Five principals from Group 1 and two from Group 2 also stressed the importance of trust within these relationships. As one principal commented:

> *[It is] difficult to effect change based on evaluation of sensitive data if the leader and host do not have trust in one another and in turn the leader builds up a team with acknowledged expertise. This is why Cyclical Reviews [are] successful. We work together and choose the leader. So we are happy to act on what we find out.*
>
> (Group 1, Principal F)

Where the team also involved members from across a learning community of schools, the trust built within the team and across the schools also appeared to be a factor that might later trigger influence. For example:

> *The team composition is good . . . builds trust. We now have staff from [the] learning community involved in [our] own and other learning-community-schools reviews. This makes planning and implementation at learning-community level good. There is carry over. They see what is happening first hand.*
>
> (Group 1, Principal F)

It can be concluded tentatively at this point that there were a range of principals' perceptions that could be grouped as human factors, comprising principals' motivations, leadership of the review team, leadership by the host principal, the leader-coach partnership, and human relationships.

As described in Chapter 5 and shown in Figures 5.2 and 5.3 data were able to be coded and grouped and this could be done iteratively as themes emerged. In addition perspectives and comments could be added to any data point to expand or clarify an emerging theme or provide context. This allowed the researcher to sift and sort through the data to identify similar phrases, patterns, and relationships between variables, as recommended by Miles and Huberman (1994:8) and allowed the factors to be further categorised into sub-factors.

Another tentative conclusion is that the principals regularly referred to openness, trust, credibility, competence, knowledge, commitment, and ownership in a way that suggested that these served as underlying values and as catalytic conditions. Again, because of the way in which data were coded, grouped, and sorted, the values that emerged were able to be attributed to specific factors.

These two tentative conclusions are shown in Figure 6.1 overleaf. The panels in Figure 6.1 are colour-coded. The organising factor is in yellow. The principals' perceptions are in blue, with their contributory sub-factors in light blue. The underlying values are provided in the orange cloud. This layout and coding will be used in Figures 6.2 to 6.4. The colour-coding of factors and underlying values remains unchanged in the rest of the study.

Figure 6.1 Mapping of Human Factors Triggering Influence

In the next section findings are presented about the evaluation factors that were seen by the principals to be triggers for later influence.

Evaluation Factors

Data-collection processes

The first evaluation factor to become apparent was the data-collection process. The seven principals from both groups who commented on the survey instrument agreed that it provided significant input to the judgements, allowed for a variance of opinions, and provided a good voice for many parents and students. They were critical, however, when the survey was not used at all or not used with a particular group—for example, not used with students—, believing that the survey would have provided much richer data. When the survey data were not ready for the team on the first day, as was the case in four of the reviews, this was also seen as a hindrance, one typical comment being, '*the earlier availability of the survey results would reduce the pressure and improve the quality of the analysis'* (Group 1, Principal B).

Principals spoke positively about those reviews where there was a wide range of interviewees, commenting that this provided a broad picture and made the data richer. In one school where focus groups of staff were used in conjunction with individual staff interviews, all three principals (Principals A, B, and K) saw this as a positive feature. They believed that the newer or more reluctant staff felt more comfortable as part of a group and therefore could provide data that might otherwise not have been collected.

There was some criticism of the interview instrument, in particular after the first review. The three principals involved in this review together with the two principals responsible for developing the interview questions for the overall process suggested that the questions needed more work as they were too open ended and did not elicit sufficient information. Principal B noted that '*when the questions aren't quite right this prevents the team from getting the right information for the school*'.

In preparation for the review host principals were meant to select and present the review team with a range of documents that supported the school's self-evaluation against the *Exemplary Practice Statements*. This component was known as the desk audit.

In cases where the desk audit was initially little more than a presentation of documents (the first three reviews), principals did not always see their relevance, three specifically commenting that the documents did not really link the rest of the review and that the documents needed to relate directly to the six domains. When adjustments were made to the desk audit following the first few pilot reviews, principals from both groups were more positive about it.

> *I got the School Development Officer to work with me on the desk audit. Much better approach—more a school self-evaluation. Worked really well. Set up the context. Helped me with time management. Understand the purpose of a desk audit much better now.*
>
> (Group 1, Principal J)

One Group 2 principal involved in a later review even went on to say that the desk audit:

> *[Is an] important factor in understanding evaluation and how to use data at school level because it means principals have to be self-critical and self-reflective. It's important because the principal can select the evidence. Some professional learning in preparing for desk audits may help principals in preparing for the review (both as team members and as principals preparing the documents). The steering committee should look to provide a set of samples of what could be included in each section (not prescriptive but giving some starting thoughts) and a couple of examples of how to provide evidence and annotate the documents.*
>
> (Group 2, Principal S)

Only six principals commented on classroom observation in terms of its being a factor that had any influence on them or required some modification. While all the principals in the working-group who developed this feature of the review thought it was a necessary and good idea at the time of development, only two of them (Principal A

and F) made a similar comment after leading a review, the rest remaining silent. One principal from Group 2 specifically commented on the benefits stating:

> *Classroom observation showed me that some staff were not following through on policy that the school had set . . . assessment and reporting, homework . . . You can only see this when you are there. When they talk about it they tend to talk as though they are doing it.*
>
> (Group 2, Principal R)

The same principals noted, however, that the process required careful judgement by the host principal and needed some rethinking because the review team tended only to see the good teachers.

PPODS

The second evaluation factor to surface was PPODS. Fifteen principals from across the two groups commented that the overall processes were sound and credible, and that these processes were a critical factor for the future use and influence of the review findings and processes. For example:

> *. . . the process [PPODS] reassured the staff more than I anticipated. One staff member said at the report session, 'I thought it would be a whitewash [but it wasn't]'. . . . The process reassured the staff more than I anticipated. It meant we could implement future directions and get staff talking about the review in meetings easily.*
>
> (Group 1, Principal A)

> *The process [PPODS] is a systematic way of collecting data. Better than methods in previous reviews, school development policy, Quality Assurance. There was a consistency in the way it was used to find themes and findings.*
>
> (Group 2, Principal S)

For many of these principals, particularly the nine in Group 2, the most useful aspects that they cited were being part of the process, seeing the wide range of data collection, and experiencing PPODS.

Of the fifteen principals who made specific comment about PPODS, twelve also thought that by further refinement the Steering Committee could strengthen the process even more by developing variations to the process for small schools, learning communities of schools, and collegiates; by increased practice; and by further development of the individual elements.

While the majority of principals liked the pocket PC, three principals felt that it was hard to use and that the same data could be collected in the same way using a laptop computer. One principal found the pocket PC '*impersonal*' (Group 2, Principal L) when recording the interviews. The researcher's observation of the teams in action confirmed that about five principals, particularly those unfamiliar with computer technology, found some of the technical aspects confusing.

However, from observation of and feedback to the researcher most principals found the pocket PCs easy to use, and enjoyed practising with them at home before the review and using them during the review. In addition, the Group 1 principals, who had taken part in previous reviews using pen-and-paper data-recording methods, commented on the increased accuracy and reliability of the data that was gained, because the information was recorded once, verified on the spot, and not subsequently lost, and the method allowed for an increased amount of data to be recorded in the time provided.

All principals recognised the need for all team members to have a common understanding of the codes and key words. Principal P from Group 2 summed up the advantage of this with the comment:

> *When all the information comes together it needs talking through as the leader might [otherwise] change the code but it is not what you heard. You need to be specific. Terrific when that discussion happens.*

Over half the principals in each group (thirteen in total) commented on the positive influence of the data analysis. From the host principals' perspective the amount of data, as measured by the number of data-point entries, was an influential factor. Principal J noted that '*it's hard to argue with overwhelming data and it's [sic] all there for everyone to see*'. Principal A commented that with the large number of data points it was self-evident to staff that this was legitimate and trustworthy data. Three team leaders commented that they would have liked everyone on the team to have the spreadsheet the night before so they could have more time to reflect, and three team-member principals who had not had a lot of experience in other types of reviews expressed some initial reservations, but were convinced when they saw it all come together by the end of the process. All team-leader and team-member principals alike found the third day of the review, when they discussed the data and wrote the report, difficult but the most influential. For example:

> *Day 3 was the hardest and best. Saw it all come together. Everyone very professional, great argument, discussion.*
>
> (Group 1, Principal E)

> *The conversation about the data was the most influential for me. I want to do this back at my school, have the conversations based on evidence, not just feelings. Process cemented that data is [sic] critical. This is the thing that still influences me and now my school two years on.*
>
> (Group 2, Principal Q)

Structures and resources

The third evaluation factor to emerge was the review structures and resources. Nearly all principals (fifteen in total) commented favourably on having the *Exemplary Practice Statements* to clarify what constituted a good school and that the statements used in the reviews were a good starting-point. One of the team leaders who had a deep understanding of the statements commented:

> *The Exemplary Practice Statements and their key role in the evaluation process need to be developed further. Once this aspect is understood the rest of the process falls into place.*
>
> (Group 1, Principal B)

Thirteen principals across the two groups indicated that this resource was one they would use in their school, even without a review, as the statements were an excellent resource for professional learning and discussion, especially with school executive and more experienced staff.

With each review, however, it became apparent to those principals from Group 1 who were also on the working party and then the Steering Committee that the *Exemplary Practice Statements* needed to be modified, and this item was raised at a number of meetings. In particular, six of the Group 1 principals believed that the statements needed to incorporate current technology practices in schools and needed to be easier to use during the Cyclical Review process. Various comments from three of the Group 2 principals, suggesting that the six domains read well but were unwieldy to use and tended to overlap, supported this latter view.

In six cases, when the timetable provided each team member with the full range of data-collection opportunities—interviews with staff, parents, and students, observations, and document analysis—the team-leader and team-member principals believed that this increased the potential for their own professional learning and also provided the host principal with more reliable data. In the two instances where the timetable was perceived as inefficient, both the team-leader and team-member principals raised this as a hindering factor. For seven of the reviews, the small number of parent interviews was also raised as a factor that might limit the validity of data, and focus-group interviews for parents were suggested as a way of overcoming this.

One of the principals from Group 1 captured most of the thoughts of principals from both groups regarding the review timetable when she stated:

> *I didn't really understand the importance of a good timetable until the review got underway. The need for time in between to reflect and download data. It's a really important feature for the school and the team. The school gets better, stronger data if the timetable works so there is the possibility that the findings will be more explicit. The team gets a richer picture and time to reflect and discuss . . . so more chance of professional learning and learning from the process. I didn't see all this until I had seen the reviews from both angles.*
>
> (Group 1, Principal G)

Despite frequent comments in the course of eight of the reviews about the intense nature of the three days, the hard work, pressure, and hectic pace, ten principals (both team-leader and team-member principals) specifically commented that the timetable was realistic and achievable and provided adequate data to make informed judgements and that the three days worked well because they did not waste any time and were not away from their schools for too long.

The majority of principals from Group 1 commented on the influence of the report from the perspective of the host principal. Eight of the nine host principals spoke positively about the report and the influence it would have on them and their schools. The features that prompted these comments were that the report was ready immediately; it was presented to the staff and handed back for use on the last day of the review; it was well received by the principal and staff because of the positive language in the achievements section as well as confirmation of directions for the school to take in the future; staff were given positive feedback; and the process was open and transparent.

The same eight host principals also believed the report format to be effective because it allowed the school community to receive the findings quickly and concisely. As Principal A, in his role as host principal, added:

> *When coupled with the data-point entries and survey results, the report provides the school with a comprehensive set of data for further investigation and future directions.*

One of the host principals, however, found the report a hindering factor for the future use of the findings and process. He wrote:

> *[There was] criticism by some staff of [the] final presentation by the team [and this] probably clouded their view of process. Also there was no final report . . . the end PowerPoint wasn't it (which in itself was fine) but mention was made of a final report pending. [The review] requires [a] stronger, clearer, final report [and a] mandatory staff briefing [is] needed.*
>
> (Group 1, Principal C)

This particular review had a different focus from the others, with the principal choosing to look at one curriculum area rather than a whole-school evaluation. With regard to the report, the team-leader principal for this review (the same principal who commented favourably on the report for his own school) commented:

> *There was also some confusion about the report and I offered to write an additional paper. In retrospect this again confirms to me that this type of evaluation needs a different type of review.*
>
> (Group 1, Principal A)

Interestingly, the principals from Group 2, while generally positive about the report-writing process, were more critical about the influence the report might have than were their colleagues from Group 1. In terms of the process and its being a factor that would influence principals, all nine commented that the process was taxing but worthwhile, especially the discussion, because they saw how all the evidence came together. They also found the process interesting and useful, because they learnt what conclusions needed to be present in the report. Five commented that the report format and framework were easy to use, simple, and easy to understand. Five also felt that the report content was accurate and that the analysis tied all evidence together to create a document that was valid. Six of them, however, also had some reservations about the explicitness of the report. For example:

> *The clarity and conclusions were not explicit enough although the school was happy. For me to want a review in my school I would want to be given more explicit future directions.*
>
> (Group 2, Principal S)

Ten principals commented on the training for the review and shared similar perceptions. They appreciated the training, especially meeting at the school and with the team prior to the review. They saw it as a great opportunity to meet fellow team members and discuss the review process and expectations. They also found that making the time available to become familiar with the technology was beneficial. Five specifically commented that training one month before the review allowed adequate time for familiarisation. They all found the most crucial aspect of the training was knowing and understanding the *Exemplary Practice Statements*.

It can be concluded tentatively at this point that there were a range of principals' perceptions that could grouped as evaluation factors: data-collection processes, PPODS (data analysis processes), and structures and resources.

Another tentative conclusion is that when the principals discussed evaluation factors they used specific criteria to describe effective processes, including quality, stakeholder voices, timeliness, relevance, participation, representativeness, promptness, comprehensiveness, and specificity. The underlying values implicit in their discussions were openness, clarity, consistency, standards, consensus, engagement, transparency, and functionality. Although in many instances implicit, the principals' discourse suggested that these values also served as catalytic conditions.

These two tentative conclusions are shown in Figure 6.2 overleaf.

Figure 6.2 Mapping of Evaluation Factors Triggering Influence

In the next section findings are presented about the context factors that that were seen by the principals to be triggers for later influence.

Context Factors

School culture

The first context factor to emerge, mainly from principals in Group 1, was school culture. All principals in Group 1 had agreed that the review team needed to

understand the school culture in order that the review report be accurate and influential. To this end they agreed to provide to the review team, as a minimum, a school context statement that would also be used at the beginning of the review report. In addition, the host principals also addressed the review team, either during the training session or at the beginning of the first day, to provide a more detailed account of the school's context and culture. Four of them commented specifically on the benefit of this. For example:

> *I liked the way the review drew on my (and the school's) self-analysis of the school. I presented this at the beginning. The self-evaluation is a really important feature and we need to build this more into the desk audit. The context presentation was good for me and the executive to prepare. Made us think what we were on about, where we were up to, what the staff were feeling.*
>
> (Group 1, Principal A)

When acting as team leaders, principals in Group 1 commented that understanding the culture and climate of the school was essential to ensure a more accurate report and to be confident that the review would have some influence beyond the principal. These same principals and one principal from Group 2 also commented that the school culture and climate needed to be one that was open and ready for a Cyclical Review. As Principal B from Group 1 noted:

> *It's important to know that staff are ready for this. Not just what the principal wants. Can't see it having much effect otherwise. Maybe in the principal's mind. That's why it's good here but maybe we are not ready for it back at my school.*

Training

Seven principals, including those on the Steering Committee, supported the idea of accredited training, giving a range of reasons. These included (1) it would be good to have a bank of reviewers in case some people needed to pull out at the last minute, (2) if the review were to influence principals, accredited training would ensure a better quality review, (3) it would be good for people, even if not used in a Cyclical Review

and good for evaluation in general, and (4) it was needed because some team members, even with training, did not really grasp how to put data together once they were collected. In summary, nearly all agreed that competencies for team members at various levels were needed because, as one principal summed it up:

> *The better you are, the more likely you and others will benefit from the review. Better preparation, better understanding, better report, better analysis . . . all those things.*
>
> (Group 1, Principal A)

Duration and timing

Principals from both groups commented on the various aspects of the duration and timing of the review. For the four principals from Group 2 who commented, the three-day length of the review was considered appropriate. It did not keep them away from their normal routine for too long and as a result they indicated that they would consider having a review in their school or being on another review. They also believed that three days was long enough, because of the accompanying processes, to gain adequate information to make informed judgements. For example:

> *Three days were fine—although action packed essential and you were right they did provide a snapshot. I think any more than that would be too much . . . too much for the school, too much for me to be away from my school.*
>
> (Group 2, Principal L)

Two principals from Group 2 commented favourably on the duration and timing of the preparation. For example:

> *I believe the amount of time we spent preparing for the review was adequate. It was close enough to the review that we remembered everything [we were taught], but enough time before going to the school to read over the documentation and prepare ourselves.*
>
> (Group 2, Principal P)

Of the twelve principals who made comments, however, ten commented on the influence exerted by the appropriate timing of the review: within the school year, the school's planning cycle, and school's stage of development. Avoiding external testing times, excursions, and the end-of-year 'winding-down' period were seen as important, with four suggesting specifically that a Cyclical Review should only occur in Terms 2 or 3, because the right time of year was essential to give the school's community the time and opportunity for reflection.

The review cycle was also viewed as important for seeing what had been achieved and what now needed to be done, with eight of the principals who commented preferring the cycle to be three to five years. Principal N from Group 2 qualified this, adding '*it depends on [the] school—size, turnover*'. Principal R from Group 2 suggested that the timing of the Cyclical Reviews should '*form part of the School Plan' as each school needs to discuss and plan . . . the best time in the planning cycle for the review*'.

Two host principals commented specifically on the timing with respect to their school's stage of development and readiness, as did their respective team-leader principals and team-member principals. For example:

> *[The] timing of review was right for my school . . . in [terms of] where the school was at, it was important to bring about needed cultural change . . . provided some immediate settling effects. Some people were still having doubts about new principal, new ideas, this changed opinion for the better—no hidden agendas.*
>
> (Group 1, Principal A as host)

> *Having led one it made me realise that timing is important. First the school needs to be well prepared for the review itself, it's not a day or two to get ready . . . needs proper preparation. But the school also needs to be ready for a Cyclical Review. Should be in relatively good shape but open and ready to accept what comes out . . . not necessarily everyone but a core of them.*
>
> (Group 1, Principal B as team-leader of above school)

Team size and composition

The third context factor to emerge was the review team's size and composition. All team-leader and team-member principals liked the size of their review teams (which varied from six to nine depending on the size of the school) and their varied composition. Their reasons included that the team was able to meld and the team size was not too intrusive, allowing at the same time for a wide range of data-collection opportunities that therefore confirmed the credibility of the review analysis. As Principal B from Group 1 explained:

> *Team formation provided a good balance of experience, expertise, and background. This is needed to get the type of report you need for future use.*

All eight host principals' comments were more concerned with the composition of the review team. They believed that having the balance of experience from school and regional staff and fresh ideas from less-experienced staff, the right staff from their own schools, staff from schools within their learning communities, and staff from other schools was a combination that provided an opportunity to receive in-depth, useful feedback about their schools' performance from an independent team.

These host principals noted that the principal's choice of which of his or her own staff would be on the review team was really important. As one explained:

> *The choice of our staff member for the review team had hidden benefits. She got a lot out of it. It built her capacity and changed her attitude. This was important for me and what I could expect to do at the school as a consequence.*
>
> (Group 1, Principal A)

Four principals from Group 1 also commented on the benefits for them of having staff from the same learning community on the team, because it helped them with planning,

activities, and getting to know the schools and students better. One principal even suggested that:

> *If staff from [your own] learning community are on the team, it keeps you honest. You are more likely to implement, because they will know.*
>
> (Group 1, Principal J)

One school invited a parent to be on the review team. The host principal (Principal E) saw this as essential as parents were part of the school community and must have this voice. She believed that it helped to cement the findings for her. The team-leader principal of this team (Principal J) commented that the parent found the data collection, coding, and sorting difficult, partly because of the educational jargon. Nevertheless, she thought that having a parent on the team was a good feature, but that it required more thought.

Review focus

The fourth factor to emerge was the focus of the Cyclical Reviews. While the concept of Cyclical Reviews had been to set up a process to examine whole-school performance against the six domains of the *Exemplary Practice Statements*, some principals wanted to use the pilot to see whether the same process could be applied to a limited number of the domains within the statements. As a result four schools examined all six domains; four schools chose between two and four domains, and one school examined all six domains, but with reference to one curriculum area only.

The two greatest concerns when schools covered all six domains, as originally intended, were, first, that there was a degree of overlap between the domains that made coding the information difficult and, second, that it was a rush to discuss and write the report when it needed six separate sections. This was noted by both the researcher as well all the principals on the teams. None the less, all principals involved

in these reviews found that it gave them a great sense of achievement to report on the school in such a comprehensive fashion. The four host principals also expressed satisfaction with the review process and report and believed that the six domains were the right focus for such an evaluation.

The greatest concern to arise when schools chose only some of the domains was that information pertaining to all the domains was raised anyway. The principals involved in the reviews where only some of the domains were evaluated provided a range of illuminating insights. The following two comments capture the essence of these.

> *The domains were limited in number to ensure enough data was collected to confidently identify trends . . . I was happy with the two as this was what the school needed to know at the time. But I admit we found out other things that I had not thought of, so maybe in retrospect all [six] would have been a better idea and we have acted on some of the others that came up anyway.*
>
> (Group 1, Principal E)

> *The school wanted the review to focus on only some of the [domains from] the Exemplary Practice Statements and that meant they were the ones that they were good at, not necessarily the ones that were critical to the school. This sort of review needs to be all the Exemplary Practice Statements.*
>
> (Group 2, Principal I)

In the school that focused on one curriculum area, rather than a whole-school evaluation, the host principal was satisfied that the review achieved what he required and believed that his review showed that a Cyclical Review could be applied to a specific aspect of the school. The team leader, on the other hand, believed that this review confirmed the view that a Cyclical Review needed to be used for whole-school evaluation.

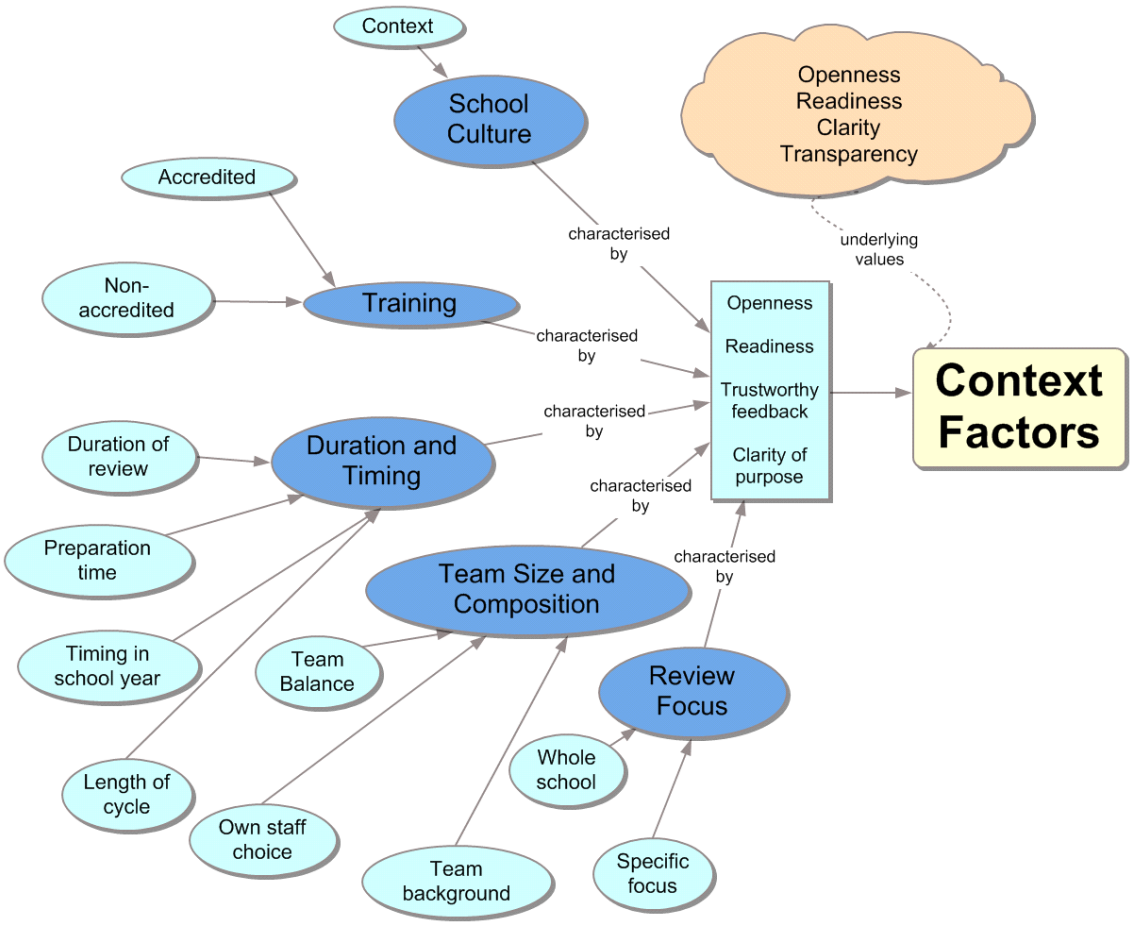
It can be concluded tentatively at this point that there were a range of principals' perceptions that could be grouped as context factors: school culture, training, duration

and timing of the review, team size and composition, and review focus. These are shown in Figure 6.3 below.

Another tentative conclusion is that when the principals discussed context factors they used specific criteria to describe effective processes, including openness, readiness, trustworthy feedback, and clarity of purpose. The underlying values implicit in their discussions were openness, readiness, clarity, and transparency. Although in many instances implicit, the principals' discourse suggested that these values also served as catalytic conditions.

These two tentative conclusions are shown in Figure 6.3 below.

Figure 6.3 Mapping of Context Factors Triggering Influence



The previous sections clarified the principals' perceptions of human, evaluation, and context factors contributing to how they were influenced. In the next section, the implications for theory, policy, and practice arising from these findings will be discussed.

IMPLICATIONS OF THE FINDINGS

To reiterate, the second research question was:

> What factors, prior experiences, and understandings contribute to the influence that the involvement in Cyclical Reviews in Western Sydney Region has had on the participating principals?

Implications of the findings related to this question are now presented in the following sections.

Implications Related to Theory

The literature (Chapter 3) examined a number of reviews of major studies into the categories and individual characteristics within them that trigger evaluation influence. Some of these studies focused on the use of results, some focused on process use, some focused on influence resulting from participatory practices, and some focused on a combination of influences.

Regardless of the focus, most of the groups and subgroups cited in these studies and described variously as categories (for example, Preskill et al., 2003), factors (for example, Alkin, 1985; King & Thompson, 1983; Leviton & Hughes, 1981), clusters (for example, Cousins & Leithwood, 1986), containing characteristics (for example, Alkin, 1985; Cousins & Leithwood, 1986), or variables (Preskill et al., 2003) were also

identified to varying extents in the current study. For example, Leviton and Hughes (1981) identified 'relevance' as one of five factors, and by Cousins and Leithwood (1986) as one of six characteristics within a cluster that leads to the use of the results. 'Relevance' could be seen as an implicit component of several of this study's identified characteristics, notably the host principal's expectations and support, timing of the review, and review focus, even though it was not defined separately in this study. 'Timeliness', identified by Cousins and Leithwood (1986) as a characteristic that leads to use of the results, was also identified as an explicit characteristic (duration and timing of the review) in this study. In this study, however, duration and timing emerged as a broader concept than the 'timeliness' characteristic of Cousins and Leithwood (1986), referring to both use of the results and process and to three aspects of duration and timing: when the review took place; the length of the review; and the timing of the training in relation to the review.

None of the sets of characteristics identified in the literature provided an exact match with the characteristics identified by the principals in the study. Therefore, while a combination of characteristics identified by various researchers could have been used to describe the characteristics identified in this study, it became more useful and succinct to assign a set of eleven characteristics that were specific to Cyclical Reviews and identified by the principals themselves. Nevertheless these characteristics had much in common with those identified through the literature.

With regard to categories, however, this study supported the groupings developed by Alkin (1985), in which multiple characteristics were grouped into three factors: human, context, and evaluation. These three factors allowed all of the identified characteristics triggering influence in all its forms to be categorised.

Thus a key finding was that there were eleven characteristics grouped into three categories that could be identified as contributing to the influence that the involvement in Cyclical Reviews in Western Sydney Region had on the participating principals. The

first category was human factors and contained the characteristics of motivation and prior experiences, leadership, host principals' expectations and support, and human relationships. The second category was evaluation factors and contained the characteristics of data-collection processes, PPODS, and structures and resources. The third category was context factors and contained the characteristics of school culture, duration and timing of the review, team size and composition, and review focus.

What did emerge as distinct from prior theory about Cyclical Reviews was the importance of underlying values acting as catalytic conditions. Their role as catalysts was most evident in the discourse of principals when they theorised about human factors that triggered influence, explicitly referring to openness, trust, credibility, competence, knowledge, commitment, and ownership.

Although more implicit in the discourse of principals, the role of underlying values as catalysts also emerged in relation to context factors and evaluation factors that triggered influence. The values related to evaluation factors included openness, clarity, consistency, standards, consensus, engagement, transparency, functionality. Finally, the context values included openness, readiness, clarity, and transparency. The implication for leadership theory about Cyclical Reviews is that these values act as threshold conditions for change.

There is a limitation that must be highlighted at this point. There is a considerable body of literature about values and leadership (for example, Dolan, Garcia, & Richley, 2006; Kraemer, 2011; Macpherson, 2014; Vilkinas & Ladyshewsky, 2011). While important, this literature and questions that emanate from it in light of the finding above are beyond the scope of this study. The finding does, however, point to further research being warranted.

The interplay between the three categories is summarised in Figure 6.4. The same iterative processes referred to above (with regard to Figures 6.1 to 6.3) were used to determine the connections between the three categories of factors.

Figure 6.4 Interplay between the Three Categories of Factors

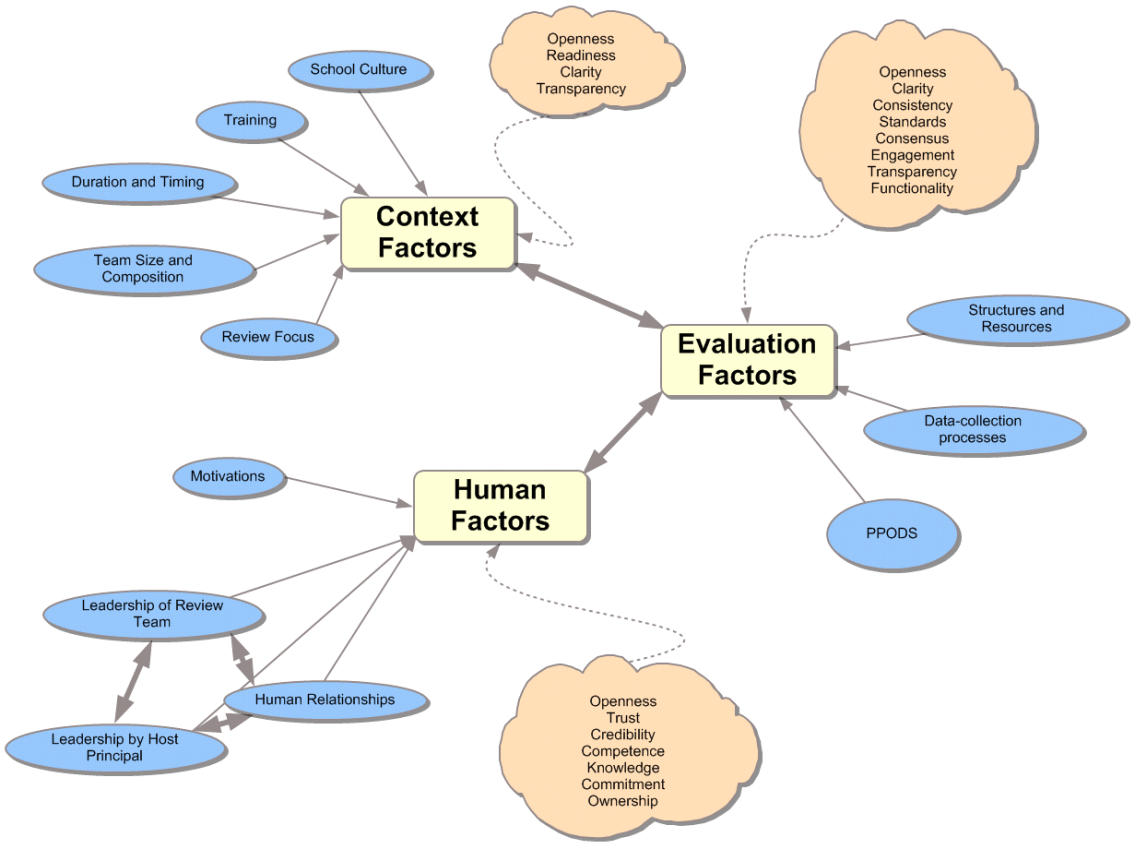


Figure 6.4 suggests that the evaluation methodology used in Cyclical Reviews is primarily independent as a process but that it is differentially affected by context and human factors in each school setting. The major variance to standardised review processes appears to be the leadership services provided to the review team, the host principal, and the partnership between the review leader and the coach.

Implications Related to Policy

Chapter 2 (context) and Chapter 3 (literature) explored a range of policy issues related to the inspectorial system, the quality assurance era, the more recent Department policies relating to school improvement and accountability, recent policies in other Australian states and territories relating to school improvement and accountability, and, in particular, the current policies of the Western Sydney Region. Direct involvement in the implementation of these policies or an awareness of them was considered to be relevant, the assumption being that the data would reveal that understanding and knowledge of, or experience in, these policies were factors that contributed to the influence involvement in the Cyclical Reviews of Western Sydney Region had on the participating principals.

This assumption appeared to be justified in part. Certainly prior experiences predisposed the principals to certain attitudes about reviews in general and Cyclical Reviews in particular. Those who reported previous positive experiences looked for similar features in the Cyclical Review process and commented favourably on those features when they were present. Similarly, the absences of features contributing to negative experiences were reported as factors that triggered influence. While most principals had experienced a range of iterations of reviews, some dating back to the era of inspections, their more recent experiences were those that most shaped their attitudes. Thus, in increasing order of influence, the principals' prior experiences of, and attitudes towards, Quality Assurance reviews, reviews under the *School Development Policy*, and the district reviews were shown to be contributing factors to the Cyclical Reviews' subsequent influence.

In saying this, three points of qualification need to be made. The first is that the personalities involved were inextricably linked with the prior experience, regardless of when the experience occurred: the human factor. This was especially apparent for the principals in Group 1, who all commented that their trust in the Regional Director,

who, as their previous superintendent, had instigated the district reviews, had cemented their trust in the current process. The human factor was not restricted to this example. There was evidence that principals who had experienced inspectors, and other 'reviewers' who had ascribed expertise in their field and who appeared knowledgeable and benevolent had also developed a certain degree of trust or openness to review processes. This points again to the threshold values identified above.

The second is that principals expressed a wariness of perceived accountability features in previous policies and models, agreeing that the developmental, inclusive, open, and transparent nature of the Cyclical Review process was a contributing factor to later subsequent influence. In particular, the development of the instruments and the process by the Group 1 principals had provided a perceived openness, transparency, and ownership that had not been available in previous review processes. Even the Group 2 principals perceived a greater degree of ownership of the process, because their colleagues had been a part of the development. Having a principal as the team leader added to the perceived developmental nature of the reviews. This reinforced the threshold values to do with evaluation.

The third is an inhibiting factor, in that principals were somewhat cautious of any new policy or process, in this case Cyclical Reviews, having been disillusioned in the past when a potentially sound policy in their terms 'just disappeared'. They had seen policies disbanded when, for example, a change in system leadership had occurred.

What these qualifications showed was that principals relied heavily on their experiences in the review implementation process and their relationships with decision-makers when forming opinions about factors they believed were necessary to make a review successful, much as King and Thompson (1983) and Alkin (1985) had proposed, rather than knowledge of the policy's development, its rationale, or its content. Almost none of the principals demonstrated an interest in, or knowledge and

understanding of, the context in which or rationale on which previous policies were based. As a result they did not refer to these factors when determining the relative merits or potential influence of the new process. In truth, some of the principals might have been too young to recall the eras of inspection or Quality Assurance. There was also a lack of interest or disinterest or both of the current political and contextual climate and current policies about reviews, accountability, and development. None, for example, mentioned policies or practices in any jurisdiction beyond New South Wales. This was somewhat surprising as the Group 1 principals in particular had taken a very active role in designing the process and it could have been expected that they would be curious about practices elsewhere; their benefits and pitfalls.

Thus a key implication was that principals' first-hand experiences in review processes and their relationships with policy-makers and those in control of the review vastly outweighed any influence exerted by policy content.

A further key implication for this study, arising from these discussions, was that principals were more willing to accept a review process and be influenced by it when they perceived the purpose to be developmental and when they could claim a degree of ownership of the process. This highlights two key dimensions of Senge's (2010) theory of schools as learning organisations when leaders take ownership of developmental evaluation, as noted above in Chapter 3.

Paradoxically, despite the expressed preference for reviews to be developmental and voluntary in nature, the principals were also willing to endorse the original vision for Cyclical Reviews, which contained elements of accountability. As described in Chapter 2, Cyclical Reviews were to provide a mechanism that would allow each school to be judged approximately every four years on whole-school practice against a set of *Exemplary Practice Statements*. There were three explicit accountability implications for policy in this vision, namely that reviews would be based on a set of standards against which all schools would be judged, that each review would judge whole-school

practice against all of the standards, and that all or nearly all schools (for example, where a school required a management review, the Cyclical Review might be postponed until the management issues were addressed) in Western Sydney Region would conduct a review over a four-year period. The acceptance of the need for a set of *Exemplary Practice Statements* and for all schools to be subject to a Cyclical Review over a period of time indicated that the principals were not averse to schools being required in time to be officially judged against predetermined standards.

Indeed, from the outset all principals expressed their satisfaction with the concept of a set of standards, in this case the *Exemplary Practice Statements*, against which schools could be judged. That they or their close peers had developed these statements no doubt increased their willingness to accept them. They also felt, however, that one set of standards was essential to ensure the ongoing quality, consistency, and credibility of reviews (factors identified, for example, by Cousins & Leithwood, 1986; Cousins & Shulha, 2006; Leviton, 2003) to gain widespread acceptance of the process and, in the immediate term, to increase the likelihood of the reports' being acted upon. Principals who had been wary of previous inspection and review processes and the motives behind them were much more accepting of a process where the standards were predetermined and where they believed the same standards would be applied in every case.

Regardless of the degree to which they later accepted and acted on their own review's findings or advocated that others should undertake a review, a key finding was that principals' attitudes towards reviews were positively influenced by the use of a predetermined and peer-developed set of standards that were applied to every Cyclical Review.

Adding to the paradox was the principals' perception that having a principal as team leader was non-threatening and supported the developmental nature of the review. At the same time they wanted this 'non-threatening team-leader principal' to, for

example, take his or her role seriously, be credible, have expertise, display leadership and technical skills, understand and rigorously analyse data, and be the right fit for the review. One even went as far as suggesting that the leader need not be a principal as long as she or he had the right skills and relationship with the host principal. In effect, principals were saying that they wanted the review to be rigorous and, provided they had trust in the leader, they were prepared to be accountable. A key finding was therefore that the position of leader of the review team was crucial and that the leader must possess credibility and the skills to conduct the review in a rigorous manner.

The issue of scope and focus of the reviews arose very early during the pilot phase, with a few of the host principals stipulating that they only wanted to look at two or three of the domains and one principal stipulating that he wanted to look at all of the domains, but only in relation to one curriculum area. That the principals were able to trial these various foci for the pilot proved beneficial, as it resulted in an almost unanimous agreement and a key finding that the process that had been developed was best suited to a whole-school evaluation using all of the domains in the *Exemplary Practice Statements*. Despite this agreement, the difficulties in addressing all six domains within the timeframe and differentiating between the domains in some instances remained an issue, with the implication being that more work needed to be done on the number of domains, the overlap between the domains, or both. The issue of becoming familiar with and using multiple codes and keywords was overcome to some degree by the overall reduction in the number of these, which occurred after the first review had taken place.

By their very name, Cyclical Reviews implied a regular evaluation that all schools would undertake. The principals, while generally positive about the pilot reviews and in favour of more, if not most, schools undertaking a Cyclical Review, were nevertheless adamant that one of the strengths of the process was its voluntary nature. Their main reason for this had two aspects. First, they believed that a voluntary review was more likely to be developmental in nature. Second, they believed this process

suited schools that were 'open' and 'ready', where staff were already implementing strategies to improve student performance and embed quality systems and were keen to look at how to improve further, and where staff, not just the principal, were willing to get involved in the process. They suggested that not all schools were at this point, and for those schools a Cyclical Review should not be undertaken in the short term. Despite this suggestion, principals were silent on how to prepare such schools for a future Cyclical Review, although it was acknowledged that different types of reviews under the *School Development Policy* were still available. Thus, a key finding was that principals perceived the voluntary nature of the process as a strength, although it was in the knowledge that schools could be reviewed 'by exception' under the *School Development Policy*, as described in Chapters 1 and 2. The key policy implication here is that voluntary Cyclical Reviews will continue to require the presence of reviews by exception to remain effective.

Despite the principals' positions about essential policy considerations for Cyclical Reviews, there remained some issues about these positions and the underlying rationale used to justify them for which the evidence collected in the course of the study did not provide findings or answers. Such policy issues included:

1. How can threshold values for effective Cyclical Reviews be embedded generically?
2. How can ownership be provided to principals to ensure that Cyclical Reviews serve as developmental evaluation?
3. How can Cyclical Reviews deliver both developmental evaluation and school-community accountability?
4. Can review by exception be retained as a fail-safe mechanism and to ensure the Cyclical Reviews remain voluntary?
5. Could the two purposes of Cyclical Reviews—that is, developmental evaluation and school-community accountability—be reconciled when promoting or implementing Cyclical Reviews across the State?
6. How could a voluntary cyclical-review process be promoted and promulgated so that most schools would willingly take part?

7. Should team leaders always be principals or should others with the prerequisite skills also be able to lead Cyclical Reviews?

It follows therefore that a key finding was that there were tensions that required reconciliation before an acceptable Cyclical Review policy for all Departmental schools could be finalised and promulgated. In particular, the tensions between ownership through participation and ownership by association, development and accountability, voluntary and mandated reviews, and team leadership only by principals or by a leader with specified skills needed consideration.

Implications Related to Practice

The various contexts, motivations, prior experiences, and relationships among the principals, coupled with plans, arrangements, communications, processes, and procedures adopted for the conduct of the Cyclical Reviews, were considered in combination under the term *practice*.

Practice, in all of its forms, was thus considered an issue in terms of the extent of its influence on the efficient and effective conduct of the Cyclical Reviews in Western Sydney Region, on the development of skills, understandings, and evaluation capacity of the principals participating in the reviews, and on the implementation of review findings and processes in the schools of the participating principals.

The initial planning and development of the data-collection processes and PPODS for Cyclical Reviews were carried out over a period of two years and involved the researcher and the Group 1 principals. There were three implications arising from this element of practice that centred on:

1. Whether the data-collection processes enabled a comprehensive range of relevant information to be collected and whether these processes, or aspects of them, were perceived by the principals as factors that would trigger future influence.
2. Whether PPODS was perceived to be a timely, efficient, and effective method of conducting the Cyclical Reviews and again whether this process was a factor that triggered ongoing positive future influence.
3. The evidence, or lack of it, suggesting that participation in the planning and development processes by the Group 1 principals was a factor that triggered ongoing positive influence, evaluation capacity building, or both.

Considering the conclusions of Cousins and Leithwood (1986) (as discussed in Chapter 3) that evaluation quality is the most important and relevance the fifth most important factor in predicting use of evaluation findings, the implication was that the more the participating principals determined the data-collection processes to be efficient, comprehensive, and relevant, leading to reliable findings, the more likely they would be to perceive these processes as factors that would trigger later influence. This proved to be the case. Data were collected by four methods: survey; interview; desk audit/document analysis; and observation. Principals' comments about these methods confirmed that it was a combination of their quality and the extent to which they were used that determined the extent to which the principals saw them as likely to contribute to influence.

In relation to the second, principals from both groups commented on the efficiency and effectiveness of the overall process, although they were critical when the data-collection processes were not used as intended or when some of the data-collection processes were not used at all.

A key finding was that the planning process proved efficient in that reviews were conducted on time and were completed following the planned procedures and using the IT resources secured for the purpose. The review teams applied themselves with

energy and conviction and the schools under review appear to have appreciated the process and the resulting recommendations for school and staff development.

SUMMARY

The factors, prior experiences, and understandings contributing to the influence that the involvement in Cyclical Reviews in Western Sydney Region had on the participating principals and that arose from the analysis of data have been identified and presented in this chapter. They were grouped into one of three categories—human, evaluation, and context factors—as provided for by the provisional theoretical model presented in Chapter 4. It was found that the Cyclical Review evaluation process varied substantially depending on school contextual factors and, in particular, on the human factors associated with leadership.

Chapter 7 will analyse the data with regard to the third and fourth of the key questions; namely, 'how does participation in Cyclical Reviews in Western Sydney influence participating principals?' and 'to what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?' Again, the data will be presented according to the provisional theoretical model.

CHAPTER 7
FINDINGS AND IMPLICATIONS ABOUT INFLUENCE

INTRODUCTION

Chapter 6 summarised the data collected in terms of the factors that had an influence on participating principals and was structured according to the provisional theoretical model developed for the study. This chapter offers a summary of the data collected in terms of the third and fourth research questions, namely 'how does participation in Cyclical Reviews in Western Sydney Region influence participating principals?' and 'to what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?' It does this by presenting an analysis of findings and a synthesis of implications of how principals perceived that they were influenced by their involvement in Cyclical Reviews. The implications for theory, policy, and practice are discussed at the end of the chapter.

SECTION 1: IMMEDIATE INFLUENCES

Results-based Influences on Principals

Principals as individuals

For Group 1 principals, two general intended influences that emerged at the individual level were the peak importance assigned to the *Exemplary Practice Statements* and the principals' improved knowledge and understanding of their staffs and schools. Six of the principals commented on how important knowing and understanding the *Exemplary Practice Statements* had now become. Principal F, for example, said that

discussing these statements '*at the beginning is really important. The results hinge on a common understanding'.*

Seven principals commented specifically on how the data and results had increased their knowledge and understanding of their staffs and schools. For example:

> *The review, especially the data and report, provided the opportunity to see how things were going and how embedded or established things were.*
>
> (Principal A)

One cognitive and affective influence to emerge was the principals' strengthened convictions about their schools' directions. Eight of the nine Group 1 principals commented on this aspect, stating that the results were not unexpected and that this had confirmed their opinion of where the school needed to head.

It was not envisaged that the review's results would have an immediate specific intended influence on the Group 2 principals and it did not.

No unintended results-based influences applicable to the Group 1 principals emerged at this level; two principals, however, noted that they were influenced by the results of the reviews that they had led. In the review that focused on a curriculum area, the team leader (Principal A) commented that '*the focus on a [curriculum area] made me consider more deeply the syllabuses and outcomes and the need to do this during Cyclical Reviews*'. Another team leader (Principal B) stated that he had '*learnt a lot*' about himself and his school '*through discussion with the [team-member] principal*' and '*results they were getting*' from this school's review.

There were, on the other hand, unintended results-based influences that emerged for the Group 2 principals. A general influence to emerge for four of them was their improved knowledge and understanding of their own staffs and schools. For example:

I thought about my school in relation to the findings so I learnt more about my school too.

(Principal K)

While this was intended for host-school principals, it was not envisaged that team-member principals would report direct influence by another school's report. This, however, also matched the unintended influence perceived by the two Group 1 principals in their role as team leaders, as noted above.

One unintended motivational influence to emerge, mainly for Group 2 principals, was the impact of the dependence on specificity in the review report. While seven principals commented that they were motivated to go back and see whether the report applied to their school, six from Group 2 and one from Group 1 also commented that the need for specificity in the review report was a motivating influence. For example:

The report needed to be more specific. I realised straight away that I would need something more specific from a review for my school to be motivated enough to have one soon.

(Principal S)

Principals as school leaders

One general intended influence emerged for the Group 1 principals. It was the principals' public acceptance of the review report's findings and commitment to their implementation. When the report was presented on the final day, the researcher observed eight of the host principals commending the report to the staff, commenting to them on its accuracy, and the credibility of the review team, and reinforcing that the report confirmed the directions the school needed to take while acknowledging and highlighting the school's strengths. Some commented on this specifically to the researcher, stating for example:

I knew from the report that I had the results that were needed for our school to improve. The staff couldn't argue because the process was open and the communication of the results was credible.

(Principal I)

While there were no unintended results-based influences that applied to the Group 1 principals as a whole, in two cases (Principals C and I) one unintended influence that was both general and motivational did emerge: the principals' justification of and motivation for supporting the review findings. In both cases, the host principals wanted to bring about certain changes, but did not have the support of the majority of their staff to do so; they each wanted the review findings to justify what they believed to be correct but unpopular decisions. Because the findings supported their position, they felt motivated to recommend and pursue the implementation of the findings.

I knew what needed to be done and the evaluation data gave me the ammunition to drive this. In saying this, however, the report itself was not well delivered and I wanted a more detailed report which we did not get. Still at the end I was satisfied that I could use the review and the report to counter many of the staff's opinions.

(Principal C)

It was not envisaged that the review's results would have an immediate specific influence on the Group 2 principals as school leaders, and this proved to be the case. Similarly, no unintended influences emerged from the data.

Principals as system leaders

One general intended influence to emerge for Group 1 principals was their improved knowledge and understanding of their learning community. When the team leader, team members, or both belonged to the same learning community as the school being reviewed, principals commented on the knowledge they had gained in relation to their learning communities. For example:

> *One of my executive was on the —— review and a teacher from —— was on ours. Both schools have gained a lot from the experience. My knowledge is better and this will help our learning community.*
>
> (Principal G)

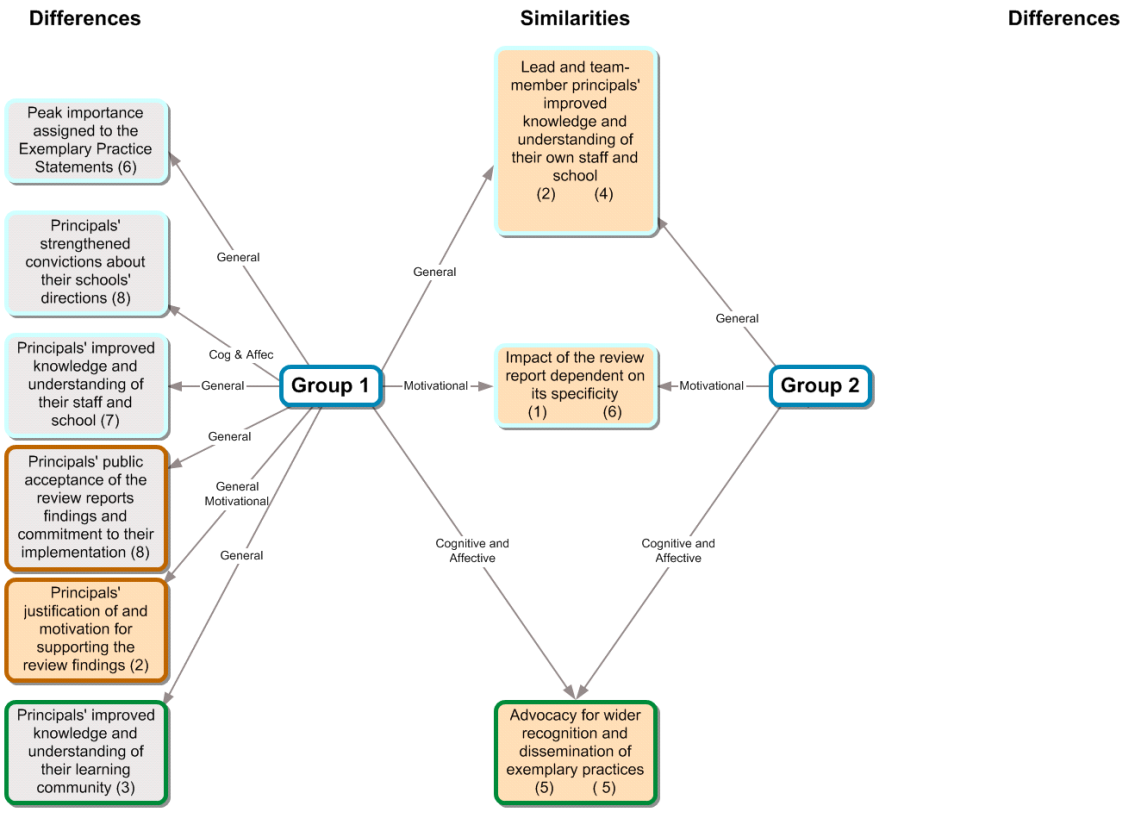
It was not envisaged that the review's results would have an immediate specific system-level influence on the Group 2 principals and it did not.

One cognitive and affective influence to emerge for five principals from each group was advocacy for wider recognition and dissemination of exemplary practices that were recorded in the review report. While it was intended that the reviews would highlight exemplary practices in schools and that publicising these practices could be used to promote and recognise Western Sydney Region schools, the actual influence on the principals was neither anticipated nor specifically intended. Typically, it was in the roles of team leader or team member that principals (as for example Principal B), referred to '*ensuring that these practices are made known across the region, . . . letting the promotions officer know about this for a good-news story, . . . and putting this on the website*'. Interestingly, but not surprisingly, in the role of host, principals in Group 1 tended to remain silent on this aspect, although all expressed pleasure when their schools' practices were recognised as exemplary.

A summary of the immediate results-based influences on participating principals is presented in Figure 7.1 below. The influences unique to Group 1 principals (that is, influences that are different from the influences on Group 2), are presented in the first column, labelled Differences. Similarly, the influences unique to Group 2 principals are presented in the third column and again labelled Differences, in this case nil. The centre column contains the influences common to both groups, and are labelled Similarities. The terms general, cognitive and affective, motivational, and behavioural are the categories of influence according to Mark and Henry (2004) as illustrated in Figure 3.2 and discussed in Chapter 3.

The cells in Figure 7.1 are colour-coded. Those with a light-blue border are the influences on principals as individuals. Those with an orange border are the influences on principals as school leaders. Those with a green border are the influences on principals as system leaders. The grey shading refers to intended influences, while the yellow shading indicates an unintended influence. The numbers of principals are provided in brackets. This layout and coding will be used in all subsequent figures (Figures 7.2 to 7.6).

Figure 7.1 Immediate Results-based Influences on Participating Principals



The implications of Figure 7.1 will be discussed at the end of Section 3, following the presentation of all influences on Group 1 and Group 2 principals.

Process-based Influences on Principals

Principals as individuals

General, cognitive and affective, motivational, and behavioural intended influences arising from the process were reported by nearly all principals in Group 1 and observed by the researcher by the end of the immediate phase. This also occurred with some of the Group 2 principals although not always to the same degree.

The first general influence to emerge was increased attention to evaluation practices. Six of the principals in Group 1 commented specifically that being involved in the working-party and then in a review had provided the opportunity for them to think more deeply about evaluation practices in general and the Cyclical Review process in particular. For example:

> *I enjoyed developing the process and spending the time thinking about evaluations. You don't give yourself that sort of time usually. There are too many other things to do. This made me do it.*
>
> (Principal G)

Only two principals from Group 2 specifically reported that participating in a Cyclical Review had led to their increased attention to evaluation practices. Principal K commented that he reflected on the process in terms of how he might use it in his own school. Principal R reported that she thought more deeply about the evaluation process—the '*what, why, and when*'—, concluding that her participation confirmed her opinion that '*you need to be ready, to know what you want out of a review before you go ahead as a principal and have one*'. Nevertheless, as with the Group 1 principals, all Group 2 principals provided many insights and constructive ideas about the individual instruments, the process as a whole, and the factors that facilitated and hindered their being influenced by Cyclical Reviews, —in this way demonstrating their increased attention to evaluation practices.

In addition, the comments and observations recorded in the previous chapter regarding the factors contributing to influence provide many examples of the principals' thinking and opinion-forming with regard to the review process.

The second general influence to emerge was increased evaluation knowledge, skills, and understanding. In particular, it was intended that principals from both groups would improve their knowledge and understanding of exemplary practice and how to evaluate it and improve their skills in setting directions for school improvement. To this end principals were asked to assess whether, as a result of their participation in their first Cyclical Review, there had been any change in their level of knowledge and understandings of and skills in (A1) discussing the *Exemplary Practice Statements*; (2) applying the *Exemplary Practice Statements* at their school; (3) planning and implementing relevant, systematic, and explicit evaluations in their current context; and (4) developing systematic and explicit targets and strategies for improvement following future evaluation processes. (See Appendix 1, Question A, p. 297.)

For the questions listed in the previous paragraph (A1–A4), eight of the nine principals in Group 1 and all nine principals in Group 2 reported increases with respect to A1, five principals from Group 1 and nine from Group 2 reported increases with respect to A2, four principals from Group 1 and eight from Group 2 reported increases with respect to A3, and four principals from Group 1 and nine from Group 2 reported increases with respect to A4.

As one principal from Group 1 further commented in relation to the first two statements:

> *I really read and thought about the Exemplary Practice Statements and worked out the key point of each entry and what that might mean for the Cyclical Review I was leading in the context of the school. As a result I feel that I understand Exemplary Practice Statements as well as anyone. [I] can apply them.*
>
> (Principal E)

Seven principals from Group 1 also commented on their acquisition of other evaluation skills and knowledge. For example, variously they perceived that they had increased their skills in developing questions and conducting interviews, understanding the significance of the desk audit, using spreadsheets, using the pocket PC, understanding and analysing data, and developing the review report. These comments reinforce and link the comments by principals cited in Chapter 6 in relation to the evaluation factors that triggered influence. Nevertheless, as expected, these same principals emphasised that they also brought skills to the process; they did not start with a *tabula rasa*.

> *I brought skills to the Cyclical Reviews. I led the survey instrument development—a combination of my existing skills and learning more about the Exemplary Practice Statements. Certainly made me think about survey design. Not sure that I am better at it yet.*
>
> (Principal F)

An analysis of annual school reports and school plans showed that six of the pilot schools relied solely on survey data to fulfil their accountability requirements with regard to evaluation. The cognitive and affective influence to emerge for these Group 1 principals was a change in their attitudes towards evaluation design. For example:

> *I have changed my opinion about doing evaluation well, with a range of ways to collect the information. It's important and gives you genuine data to work on.*
>
> (Principal I)

Principal A, representative of the others who were already well versed in and employed a range of evaluation approaches, also commented that Cyclical Reviews had reinforced his opinion that every evaluation needed to be specifically designed to meet the requirements at the time. He suggested that this may involve a survey, but it was more likely to be a combination of data-collection devices from as many sources as possible, bearing in mind that it would always be a bit of a juggling act between needs, time, and resources.

As with Group 1 principals, this same cognitive and affective influence emerged. Seven of the Group 2 principals commented similarly; for example:

> *I have been working on other ways to do evaluation and this showed me that a survey isn't the only way to go.*
>
> (Principal O)

One intended motivational influence emerged for principals in both groups: seven from Group 1 and five from Group 2. It was the personal desire by principals to continue to learn. Principals variously reported that they were motivated by being involved in a real professional learning activity, doing a real job that is needed, and being helped along the way by people who know what they are doing. As a principal from each group summed it up:

> *You can always learn. That's the real motivation. I gained more knowledge of the process, the timing, what was required of a leader in a Cyclical Review. Although I knew about it and had worked on the concept [and] had led many other reviews, it is not the same as actually leading one [of these].*
>
> (Principal A)

> *I do not get any rewards for participating. It's more intrinsic . . . what I can take back to my school. [The] best part for me was being in a professional learning activity that reinforced concepts of effective evaluation.*
>
> (Principal S)

One behavioural influence to emerge was the demonstration of newly learnt skills. The researcher observed six principals in Group 1 gain and demonstrate competence in all aspects of the Cyclical Review process and the others gain confidence and competence in some aspects of the process. As one principal commented:

> *I felt the pressure. I thought I could do more in the way of leading but it was a bit daunting. However, I got better with each day. I would like to do it again. I learnt a lot.*
>
> (Principal G)

While observation of the Group 2 principals during the training and conduct of each review showed that they had acquired new skills and they were putting these new skills into practice, none of the principals specifically commented on this as an influence. Nevertheless, the performance of new skills—collecting, coding, and analysing data, using spreadsheets, and interviewing techniques—was demonstrated to the researcher by seven of the principals. One principal (Principal N) did, however, feel that he had not been able to acquire and therefore demonstrate the new skills required. He found that for him the short preparation time was a disadvantage, not providing him with the time to learn the new skills required and leading him to feel that he did not contribute as well as he could have.

The decision to pilot the reviews in schools where the Regional Director had previously been the superintendent was made because it was presumed that principals in these schools would be more open to the process and more willing to volunteer, owing to their involvement in reviews in 2003. What emerged as an unintended motivational influence was the sway that the Regional Director had on the Group 1 principals. It was not just the experience in earlier reviews that influenced the principals to participate; it was, as put by six principals, the Regional Director's '*vision*', '*personal style*', and '*leadership*'. As Principal H summed up:

> —— *set the standard. The focus is on learning. An unrelenting focus on learning. I need to do that too.*
>
> (Principal H)

While it was intended that all principals would further develop skills in evaluation, it was also anticipated that Group 1 principals would feel comfortable with implementing the process and be prepared to lead and host future reviews. An unintended, though not widespread, behavioural influence to emerge, therefore, was the anxiety expressed by three principals in implementing and performing new skills. Furthermore, it was unanticipated. For example:

> *I would need more practice or more assistance before I could run a good evaluation on a different topic or at least in a different way. You know . . . if you have something you want to find out, how do you best get the answer and data?*
>
> (Principal J)

No unintended influences emerged for Group 2 principals and none of them mentioned the Regional Director's leadership or influence as a motivation.

Principals as school leaders

It was intended that principals in Group 1 would be satisfied with the conduct of the review and its findings and therefore speak positively about the process and outcomes during the report presentation. In nearly every case (eight of the nine) this occurred, and consequently a cognitive and affective influence to emerge was advocacy within the school for the Cyclical Review process. There were no unintended influences observed or reported. It was not envisaged that the review's processes would have an immediate specific influence on the Group 2 principals and no unintended influences were apparent from the analysis of the data.

Principals as system leaders

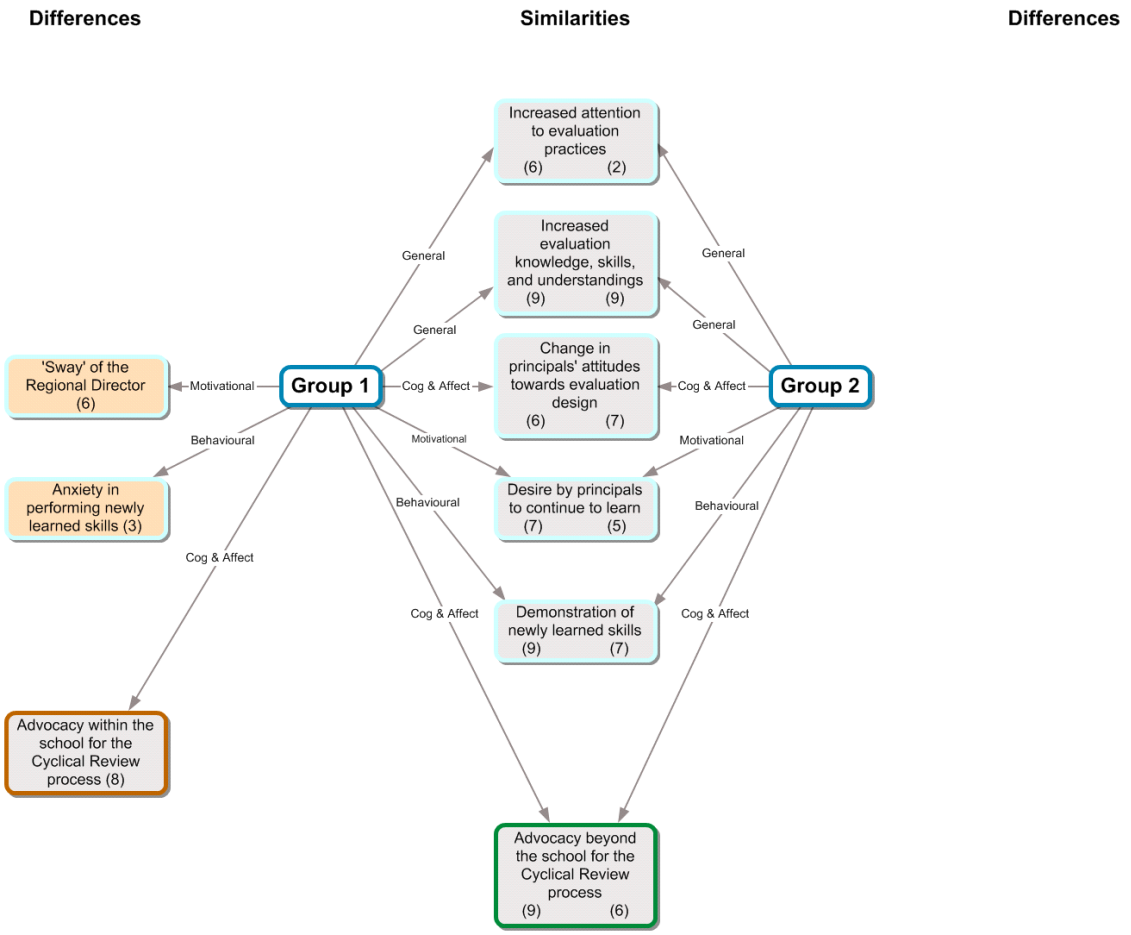
One intended cognitive and affective influence to emerge was that of advocacy beyond the school for the Cyclical Review process. Every team-leader principal commended the process to the host school's staff during the presentation of the review report. This added external support to the host principal's view that the process was a valid one. Principals at the working-group meetings also voiced their support to the group and the Regional Director. One reason for this advocacy, as described by Principal A, was that hosting or leading a review had shown them that the process built the capacity of principals as team leaders; that they had seen this occur with the leader in their school and experienced it as a leader themselves, not just as a leader of reviews but also in the

way they worked with teams at school. Another reason, as described by Principal G, was that the process was a great opportunity [for principals] to see at first hand how another school was running. A third reason, as described by Principal F, was that it was a good opportunity for each learning community to learn more about themselves. Advocacy for the process beyond the school was also observed by six principals in Group 2. Where team-member principals spoke at the presentation of the review report to the staff of the school being reviewed, they commented on the excellent process and how they were impressed with it, reassuring staff that the results were a collective and were based on valid data.

No unintended influences were apparent from the analysis of the data for either group.

A summary of the immediate process-based influences on participating principals is presented in Figure 7.2 overleaf.

Figure 7.2 Immediate Process-based Influences on Participating Principals



SECTION 2: END-OF-CYCLE INFLUENCES

Results-based Influences on Principals

Principals as individuals

One general, one cognitive and affective, one behavioural, and one motivational intended influence emerged for Group 1 principals, and all these influences were interrelated. The general influence to emerge was the reflection by principals on their schools' achievements and future expectations. Six principals reported that in the few

weeks following the review they thought a lot about the positive aspects of the report: the achievements, the most positive survey responses, and the domain summaries; and they thought about their future expectations for these areas. In their role of leading the reviews these principals also commented that they thought a lot about the good practice they had observed.

A related cognitive and affective influence to emerge for six Group 1 principals at this level was that one or more of the findings in their review reports increased in importance in those principals' eyes. As one principal stated:

> *The review raised where we were not doing some things. These took on a greater importance. I used this . . . thought about what I now needed to do . . . if I needed to do anything.*
>
> (Principal A)

While the areas that took on a greater importance differed for each review—for example, '*addressing higher order thinking skills*' (Principal G), '*matching programming to the syllabuses*' (Principal A), and '*how the executive is modelling, mentoring, and encouraging staff*' (Principal F)—, they directly linked to the areas listed for future direction in each review report. The same six principals, however, stressed that they also paid attention to every part of the report and thought about the priorities for action.

With this increase in importance of one or more areas, five Group 1 principals perceived a behavioural influence; they began to change their own practice, most often by talking more about an area that had become important to them as a result of hosting or leading a review or by putting it on the [staff or executive meeting] agenda for discussion. The latter was verified in the meeting minutes of the schools concerned.

A further related influence was motivational: the increased impact of the report as a result of its being developed against a set of standards. Six Group 1 principals reported

that judging their schools' performances against their own agreed-to standards (*Exemplary Practice Statements)* made the report findings more valid, more genuine, or more believable, and therefore made the need to implement the future directions more compelling.

> *I think because the review data is judged against a set of standards, the results are more likely to be accepted, at least that's the way I see it. It is good to have standards to aspire to. Our report had more guts because we developed the standards. It seems more believable that way. So I'm more inclined to believe and implement. They are not just one team's view of the world. The report is judged against what we believe are high standards.*
>
> (Principal H)

With similar comments to their Group 1 colleagues, five Group 2 principals perceived that, in the weeks following the review, the results provided a stimulus for them to reflect on good teaching practice, quality teaching and learning, outstanding school leadership, ideas for programming, excellent school management and planning, or a combination of these. As one principal commented:

> *Once back at my school, I thought about the review report. Each of the achievements—good teaching practice, outstanding school leadership—became important in my mind as I examined them, thought about them some more, and determined what my expectations were for my school.*
>
> (Principal S)

Principals from Group 2, unlike their Group 1 colleagues, were silent about any motivational influences. One possible reason for this is that they did not have the same involvement in the development of or discussion about the *Exemplary Practice Statements*, and although they expressed their support for their content and their use as a set of standards, they did not necessarily see them as their '*own agreed-to standards*' (Principal A).

Although it was intended that all principals would gain insights into schools' strengths and perceived areas for development and that as a result they would think more

deeply about issues that emerged, it was intended that only host principals would then give some issues greater attention and heightened importance, and make decisions based on this. Acting in a similar manner on the basis of another school's results was neither a foreseen nor an intended outcome. Nevertheless, there was one instance where a team-leader principal reported these same influences (increased importance placed on one of the findings, and a change in his own practice) occurring from the results of the review that he led.

> *I used what I had learnt about the school's KLA [key learning area] to consider what we were doing at our school. This influenced me in some of my KLA-based decisions and beliefs back at my school.*
>
> (Principal A)

While reported from only one principal in Group 1, this is significant, because these influences were also reported by four principals in Group 2. For example:

> *There were a range of future directions connected with PBL [Positive Behaviour for Learning]. It was a program that I was embarking on and it made PBL much more important for me. Some of the strategies for my school come directly from the review.*
>
> (Principal P)

Not every principal, however, felt this influence; one, in particular, commented:

> *I didn't take any of the findings back. They were not applicable to my school, [but I] enjoyed the experience.*
>
> (Principal L)

Principals as school leaders

General, cognitive and affective, motivational, and behavioural intended influences emerged for Group 1 principals in the months following each review. In fact this period represents the most time and highest level of activity and influence on this group throughout the study.

Both general influences and cognitive and affective influences were the first to emerge. They emerged as the two roles played by principals (namely, agents of change and advocates for the findings) when addressing the review data, findings, and report with their executive, staff, and community. All host principals reported that within a few weeks of the review's completion they '*talked about the report*' with their executive, '*presented the report again*' at a staff meeting, '*presented the report*' at a parent meeting, and '*made reference to the review*' or '*listed the main findings*' in school newsletters or on websites, two adding specifically that:

> *If you want to implement the changes, . . . confirm the good things, and . . . generally make use of the review, . . . this is what you need to do.*
>
> (Principals A and F)

To a much lesser extent these influences were evident for five of the Group 2 principals, who stated that they had reported the good or exemplary practice that they had observed to their executive teams on return to their schools. Two principals (Principals Q and S) added that they had listed specific items on their executive or staff meeting agendas, and minutes of the meetings confirmed this. For example, Principal Q's school had the 'Cyclical Review—report' listed as an item on the executive meeting agenda in the week following the participation of the principal as a team member and the minutes referred to a discussion about the good features of homework, assessment and reporting, and induction observed during the review.

As mentioned previously, in one school the team leader (Principal J) came back in the following weeks to co-present the report to the community and to discuss the report in more detail with the staff. In this instance the host principal (Principal E) specifically commented on the positive influence this had on the acceptance of the report and the desire to move forward, both for herself and for the wider school community. She added, '*it made my work that much easier and I was also able to probe the findings more thoroughly*'.

The next intended influence to emerge for Group 1 principals was motivational. Leading and hosting a review confirmed in most principals' minds the benefits of having the results based on a set of their own agreed-to standards, as discussed above. In turn, seven principals, when addressing the executive and staff in the weeks following the review, compared the school's performance, as described in the review report, with the agreed standards outlined in the *Exemplary Practice Statements*. In this way principals conveyed to their staffs that the *Exemplary Practice Statements* provided the standards that the school should aspire to. Nevertheless, three were quick to add that while they encouraged their executive and more senior staff, who had supervisory and planning responsibilities, to know, understand, and apply the *Exemplary Practice Statements*, this was not as necessary for less experienced or beginning teachers, who still needed to concentrate on the *Quality Teaching* framework to get their classroom teaching right first.

The reviews for seven of the nine schools took place between September and December. This is also the time of the school year when the year's achievements are discussed, evaluations are written, and the Department requires schools to develop their annual school reports and develop or modify their school plans. Owing to the simultaneous timing of these seven reviews and the Department's reporting and accountability requirements, three further influences emerged within the end-of-cycle timeframe: one motivational, one behavioural, and one general. The Department's imperatives provided the motivational influence for principals to use the review results in their school's planning and reporting processes and documents. By doing so, principals were further influenced to incorporate the review findings into their following year's school plans (in essence a change in or strengthening of school policy), having first been influenced by their own and their staffs' responses to the findings, in the form of new or revised targets, strategies, and outcomes (in essence the drafting of policy and the setting of standards). As one principal commented:

> *I would have asked my executive to discuss the findings with their teams and look at what we needed to do anyway, but maybe we would have taken a bit more time . . . but*

the timing of the review coincided with the our end-of year reporting and development of our next year's plan . . . so we did it straight away.

(Principal H)

Document analysis also confirmed that each of these seven schools reported on its Cyclical Review and used the review report for the required evaluations in its annual school reports. Further each school also incorporated the review findings in its 2008 school plan. For example, in Principal F's school the Cyclical review report recommended the 'expansion of student-centred learning to include all students, including English as a Second Language (ESL) students'. Its 2008 school plan included a focus on 'expanding student-centred activities and explicit teaching strategies for ESL students in the mainstream classroom', with specific strategies to achieve this and targets by which to measure success.

The final intended influence to emerge for Group 1 principals was a behavioural one and one that began to emerge towards the conclusion of the end-of-cycle timeframe. It was the changes in practice that occurred across the school. For example, in the school mentioned above, the principal noted that the changes evident in the plans were also being put into action by the executive and staff. Teams engaged in discussions about student-centred learning and tried a range of explicit teaching strategies. Another commented that:

We first developed an implementation plan, with teams taking responsibility for various findings in the report and before the end of the year each [team] had reported progress with implementation. The review findings highlighted that more emphasis is required on the recognition of learning outcomes in classrooms. Results confirm this too. We have included this in the annual school report. It confirms that action has already begun in this area with changes to and an increase in the number of student academic excellence and achievement awards.

(Principal A)

In contrast to the implementation processes described above, two Group 1 principals reported that they did not need much more discussion once the review report was

tabled. As Principal I noted, '*it confirmed what I wanted to do, so we took the review findings and implemented them*'. The influence of the results on these two principals as school leaders was one of justification for their own views. This was certainly not intended. For Group 2 principals there were no unintended influences observed or reported.

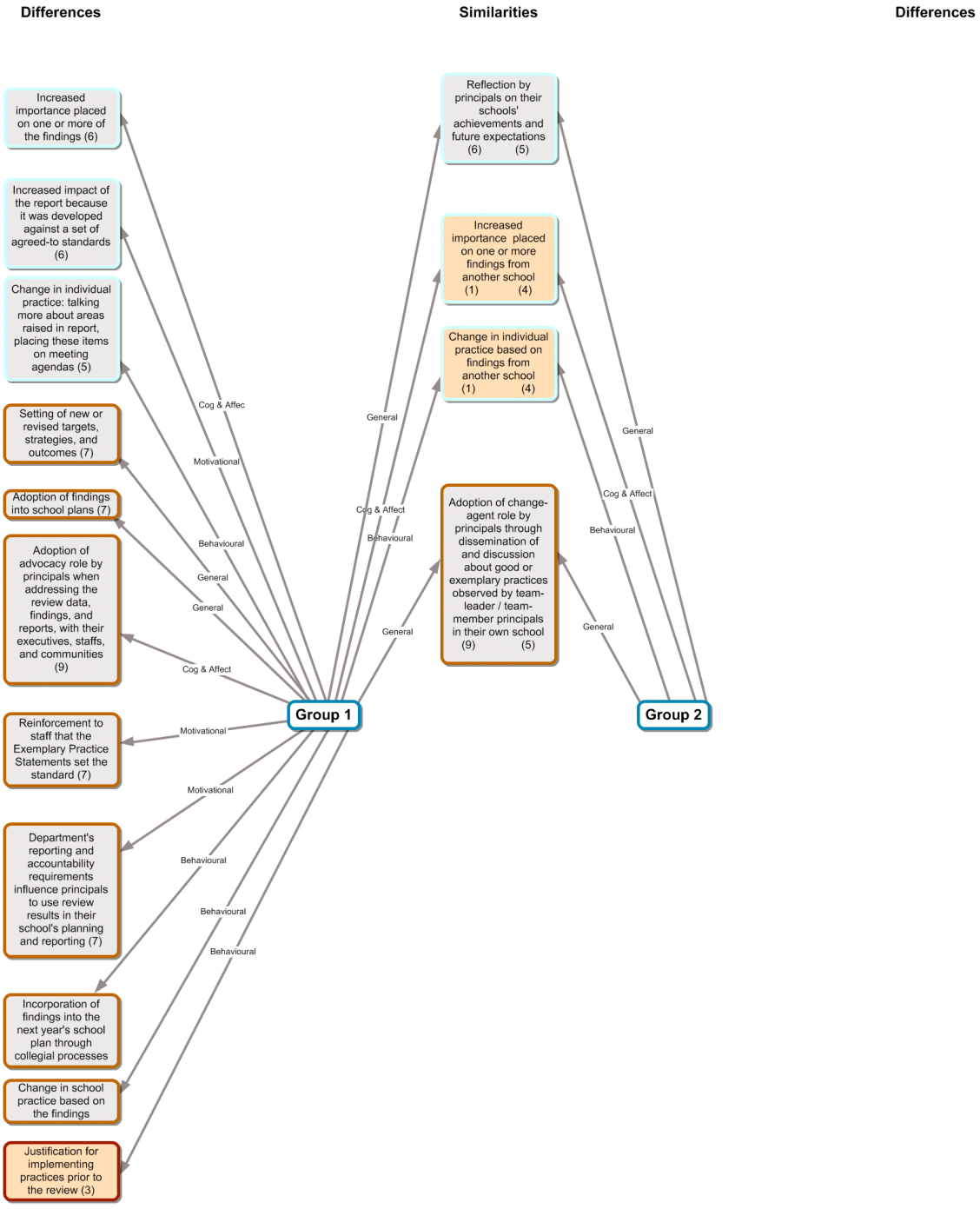
Principals as system leaders

It was intended that good or exemplary practice would not only be recognised by the Cyclical Review process, but that principals from both groups who had taken part in reviews would make these practices more widely known; for example, by commending these practices to their colleagues. It was also envisaged that the region would recognise exemplary practices and promote them. While most principals commented at the time of the review that particular practices were excellent and should be acknowledged publicly, and nearly all added that they had discussed at their own schools the good things they had seen, none reported that they had acknowledged such practices more widely. The working-party also discussed the need to find a suitable avenue for the promotion of exemplary practice, but deferred any further decision until the establishment of a regional website. Thus the intended general influence of advocacy for and wide dissemination of exemplary practice did not occur. To his credit, the Regional Director did acknowledge in his weekly newsletter the reviews as they occurred and at principals' meetings, and commented that they were highlighting the excellent work occurring in the region's schools.

There were no unintended influences observed or reported for either group.

A summary of the end-of-cycle results-based influences on participating principals is presented in Figure 7.3 overleaf.

Figure 7.3 End-of-Cycle Results-based Influences on Participating Principals



Process-based Influence on Principals

Principals as individuals

For Group 1 principals, general, cognitive and affective, motivational, and behavioural influences continued to emerge, although not in all instances at the level expected.

The general influence of increased attention to evaluation practices that emerged in the immediate phase continued for a few months after most principals had either hosted or led a Cyclical Review or done both and therefore continued into the end-of-cycle timeframe. During this time their attention was drawn specifically to Cyclical Review processes, rather than evaluation practices in general. The main reason for this, as observed by the researcher, was that the principals remained actively involved in the working-party, where the processes were discussed and further refined. As a result they would telephone, email, and occasionally meet face-to-face with the researcher, with comments such as:

> *I have been thinking about the interview questions and the codes. I have looked at asking other sorts of questions. I'll send you some examples. Also I have been thinking about the codes and I think we could improve them. My thinking is there are too many codes we should limit them.*
>
> (Principal B)

The researcher observed that their comments tended to be about the process in the development of which they had had the most involvement or the one that they had found the most difficult to manage as a team leader. The most frequent comments were about the *Exemplary Practice Statements*, with seven principals suggesting that they needed more work or revision to make them manageable. This was also the area in which the researcher, as coach, had been engaged with the team-leader principal for the longest time during the review itself; in particular in discussion about the sort of

evidence that related to each domain and whether data belonged to one domain or another.

Despite regular reflection on evaluation practices, only three principals reported that their evaluation skills, knowledge, and understanding continued to increase, even though this was an intended general influence during the timeframe. Of these, two commented that their knowledge of the *Exemplary Practice Statements* continued to expand because they continued to apply the statements to their daily work, and two commented that, since the review, they had developed better skills in linking planning, reporting, and review outcomes, owing to the logical and analytical process that they had participated in.

One influence to emerge that was both cognitive and affective and motivational for six principals was the strengthening of their positive attitudes towards the Cyclical Review process. As one principal at a working-group meeting summed it up:

> *Looking back it is the process that makes the reviews worthwhile. I agree we need to modify some components . . . make them easier to use . . . and step up some parts of the training, but the process makes the reviews work. The staff accept the result. So did I, because it really confirmed what we knew was good or needed to be done. But it also helps the staff to open up, to discuss things more, to question more. To start with, I was happy to go along with the idea and be in the pilot. It's more than that now. It's a good process that we have developed and it helps us develop. It's not just what someone else wants or says is what we should be aiming for.*
>
> (Principal C)

For four principals a behavioural influence emerged. It was a change in the way they approached evaluation. These principals commented that they now approached evaluation more logically and more systematically and were more inclined to look for evidence. Although this was an intended influence, the other five principals did not report any change in their practice in this regard. One reason for this lack of change might have been the timing of the Cyclical Reviews within the school year. The end-of-cycle phase coincided with the final term of the year and the first term of the following

school year. With the Cyclical Reviews fulfilling many of the evaluation requirements for participating schools at the end of the year and with very little formal evaluation typically taking place at the beginning of a school year, the opportunity to demonstrate a change of practice was not readily presented.

The intended influences for the Group 2 principals were somewhat different. A range of general, cognitive and affective, motivational, and behavioural influences emerged, although some aspects of these influences applied to individual principals rather than the principals as a group.

The general influence to emerge for six of the principals was that they became more reflective about the Cyclical Review process. They commented that they thought about the process, how or whether they would use it in their school, and what they would expect such a process to achieve in their own school. Three of the six also commented on the ongoing increased attention that they now paid to evaluation practices in general. Principal S, for example, stated that:

> *Gathering data from different sources made me think about what good teaching looks like and how I would know. I asked the executive to apply the same questions and processes. What does it look like? How do you know? I asked them to use same approach with their teams.*
>
> (Principal S)

Three principals also commented on their increased knowledge of quality practices and what was happening elsewhere.

A further general influence to emerge was that of the increased importance accorded to particular features of the review process. Nevertheless, principals had different views about the relative importance particular features. For example, for two principals the most important feature of the review was team work. For another:

> *[I] realised that knowledge of key words and the exemplary statements and their relationship to coding and data analysis needs to be established to provide a framework for the process is essential.*
>
> (Principal Q)

For another principal it was the desk audit that was deemed the most important feature because:

> *You need to know your school, to have thought about your school, and analysed it first for the review to have the best chance of bringing about any changes that are needed.*
>
> (Principal R)

For Group 2 principals a cognitive and affective influence that emerged was the cementing of positive attitudes towards the Cyclical Review process. Seven said that they were satisfied with the review, and that they would recommend the process to others. One principal (Principal Q) went further, stating that he was convinced that his own school would benefit from a Cyclical Review, and another (Principal K) was convinced that a form of Cyclical Review would be worthwhile in his school. Two reported an even stronger conviction about the positive influence of teamwork in effecting change. Three principals reported being even more focused on collecting and using data from a range of sources and two others reported that they now felt more positive about Cyclical Reviews and had more confidence to engage in the Cyclical Review process.

Two principals' comments showed a motivational influence. Both perceived that the review had provided an avenue to fulfil their own professional-learning goals. One, (Principal S) who was an experienced principal, believed that the process had helped him to find out more about the '*how of evaluation*'. The other (Principal P), fairly new to the principal's role, believed the process had provided him with the opportunity to see '*experienced principals in action*', which was good for him '*to develop as a leader*'.

Two principals reported and demonstrated a behavioural influence. Both applied the *Exemplary Practice Statements* to think about and review in their minds the strengths of

their schools and areas for development. Both commented that without the review they would not have done this, or at least not in this way. One principal, however, reported that:

> *Although I had good intentions to use the Exemplary Practice Statements and analyse my school after the review, I returned to school, continued in my normal role, and did not implement any changes. That may be because I am still getting used to being a principal.*
>
> (Principal P)

One behavioural unintended influence to emerge for three principals in Group 1 and one in Group 2 was the ongoing personal use of pocket PCs. A few purchased their own pocket PC as a result of learning to use them and seeing the wider possibilities for their use during the review process. One principal (Principal B) commented that he now used his own pocket PC for all note-taking and diary entries. While it was intended that the Cyclical Review instruments and procedures would be used in future school-evaluation practices, the appeal of pocket PCs for other applications was not foreseen. No other unintended influences emerged from the analysis of the data.

Principals as school leaders

Some general, cognitive and affective, and behavioural intended influences emerged for individual principals in Group 1, with one cognitive and affective influence and one behavioural influence applying to the group as a whole.

Seven principals reported, and meeting agendas and minutes confirmed, that in the months following the review these principals had become aware of strengths in their staffs that they were previously unaware of or that they saw new strengths and attitudes developing in their staffs. Often this was as a result of increased staff discussion and open but healthy debate. As one principal stated:

> *The level of professional discussion increased among staff along with the number of teachers actually voicing their opinions. This previously was the domain of only a few in the staffroom. I put this down to the process. I saw different abilities and skills in staff as a result.*
>
> (Principal I)

Another principal (Principal A) reported that he used the '*review process to assist in bringing about a cultural change at the school*': a general influence. As in a previous example, he also described the cognitive and affective influence: how attitudes towards the staff member who was on the review team had become more positive, because he and others '*became aware of some of her strengths*'. Further, he reported that the process had helped to bring about a change in school practice: a behavioural influence that was either stated or implied by seven of the other host principals.

> *[The review] had a positive impact on the way we do business — facts, data, listening to each other, openness, transparency. The Cyclical Review model is far superior to other models, it builds capacity. We took it back to other operations in the school.*
>
> (Principal A)

One principal (Principal B), however, reported that the intended general influence of bringing about cultural change did not happen. Having led a review, he tried to persuade his staff, by (amongst other things) demonstrating the pocket PCs and PPODS, to host a review at his own school, but reported that the staff were still not convinced about the merits of the process and as a result there was an executive decision not to proceed with one.

Although the evaluation sections in the 2007 annual school reports for Group 1 schools showed, in general, an improvement in the schools' capacity to conduct and report on management and curriculum evaluations, this occurred because they all reported the Cyclical Review findings and processes. It was therefore not possible to determine whether there had been a change in practice that indicated increased evaluation capacity, although this was an intended influence. Nevertheless, during this time, one principal (Principal F) decided to apply PPODS to evaluate another issue that had been

highlighted through her school's Cyclical Review. She and other staff who had already taken part in a Cyclical Review taught others, including students, how to plan the evaluation and collect the data, and, as a result, carried out what they determined to be a very successful and valid evaluation.

Again, for the Group 2 principals the intended influences were somewhat different. It was intended that in the weeks following their participation, principals would return to their schools and speak positively about the review process. In nearly every case (eight of the nine) this occurred, and consequently a cognitive and affective influence to emerge was advocacy within principals' own schools for the Cyclical Review process. It was also intended that these principals would introduce parts of the process to their staff, thereby acting as a change agent. This general influence was reported by four of the nine principals. For example, three reported introducing some or all of the *Exemplary Practice Statements* to the staff, and three reported using the Cyclical Review processes with their executive to examine data.

Seven of the principals also reported a behavioural influence, now believing that they were more likely to look for evidence to support proposed changes or to verify improvements and to expect their executive to do the same. As one principal summed it up:

> *I asked the executive to apply the same sort of questions that were used in the review. Such as, What does it look like? How do you know? . . . and to use same approach with their team. Then we discussed the evidence that they collected.*
>
> (Principal S)

It was further intended that participation as a team member would motivate principals to gain the support of their staffs and commit their own schools to a Cyclical Review. Only one principal, however, was influenced in this way.

No unintended influences emerged from the data for either group.

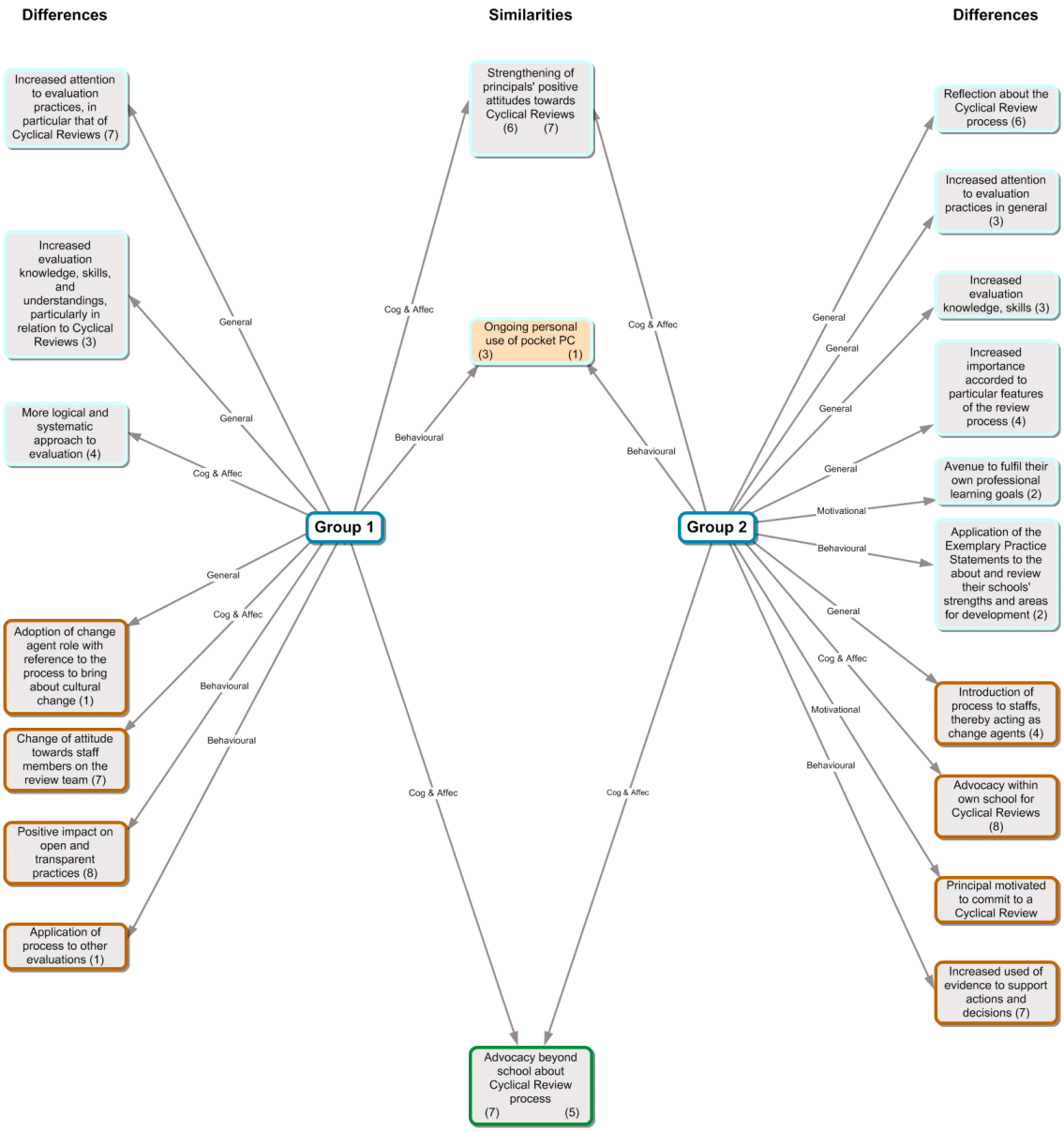
Principals as system leaders

The only intended influence to emerge was the cognitive and affective influence of advocacy beyond the school for the Cyclical Review process. Seven principals in Group 1 reported speaking positively about the process to at least one colleague beyond the pilot group. Five principals from Group 2 also reported that they had recommended the Cyclical Review process to some of their colleagues, although the majority of these recommended being on a review team as a worthwhile experience rather than having a Cyclical Review in one's own school.

There were no unintended influences observed or reported for either group.

A summary of the end-of-cycle process-based influences on participating principals is presented in Figure 7.4 overleaf.

Figure 7.4 End-of-Cycle Process-based Influences on Participating Principals



SECTION 3: LONG-TERM INFLUENCES

Results-based Influences on Principals

Principals as individuals

One cognitive and affective intended influence was observed for the Group 1 principals: that of a shift in attitude. Six of the nine principals were observed over time to have gained confidence in talking about their schools' strengths and were very open about their own schools' findings and the work of their schools as a result of the review. As Principal G commented:

> *I am happier now to talk about my school and where we are up to. The report confirmed my beliefs and gave me legitimacy to talk about our school. It's not so much that I changed or found out new things about my school its more that my attitude towards things changed and I felt much more comfortable affirming what I now know is happening and why.*

It was not envisaged that the review's results would have a long-term specific influence on the Group 2 principals, and this proved to be the case.

There were no unintended influences observed or reported for either group.

Principals as school leaders

Two intended influences based on the results, one general and one behavioural, were reported by the principals in Group 1 and were also observed by the researcher during the long-term phase of data collection.

The main general influence to occur was that of policy development. At the end of 2007 schools were required to update their school plans for 2008 and towards the end of 2008 schools were required to develop their next three-year strategic plans. In every case of a school that had undergone a Cyclical Review, the future directions from the review were evident in each of their plans. For example, in the case of Principal G's school, the review report contained a recommendation that professional learning needed to include further work on quality teaching, include classroom observation of peers as a strategy, and develop teachers' assessment and reporting knowledge. Throughout the school's strategic plan for 2009–11 there was an emphasis on quality teaching with a focus on assessment and on teaching and learning practices. Strategies included in the plan included collegial visits, lesson observations, and professional discussion of teachers' learning from these strategies using the quality teaching framework. The principal also specifically commented on this, reporting:

> *The future directions from the Cyclical Review continue to form the basis for direction of the school planning and evaluation process for the 2009–11 school plan.*
>
> (Principal G)

The behavioural influence to emerge was the principals' belief that the adoption of changes as the result of the review had now become embedded in school culture and practice. Possibly naïvely, it was envisaged that principals would continue to be directly influenced by the findings and the report and to refer to them when discussing, analysing, or evaluating school planning, strategies, and directions with their staffs, even after one or two years. This was not the case. Although all Group 1 principals were to some extent conscious of the ongoing influence of the review, seven specifically noted that changes had now become embedded into school practice and culture so that the review itself no longer needed to be referred to directly. For example:

> *The results and directions seem embedded . . . embedded into the school plan. The school has moved beyond the need to reflect on this particular review. It's the plan that has now become the direction of the school. So the review is not referred to directly but the actions are.*
>
> (Principal A)

Indeed the general consensus was that principals used the review report and findings directly at school level for about eighteen months, by which time the findings were embedded in school practice, or the school had moved beyond the future directions contained in the report.

For Group 2, it was intended that principals would gain valuable professional learning from being on a review team and take back ideas, initiatives, or strategies to their own schools for discussion and implementation. In four cases this was reported to have occurred. For example:

> *I took back ideas about the use of technology to engage students in their learning and have implemented some of these in consultation with my executive team and staff.*
>
> (Principal O)

Consequently a behavioural influence to emerge was a collaborative change in practice.

There were no unintended influences observed or reported for either group.

Principals as system leaders

It was intended that good practice observed and reported during a Cyclical Review would be used by the Group 1 principals (and others) on the Steering Committee to promote not only these practices across the region but also to promote Cyclical Reviews. This motivational influence did not occur.

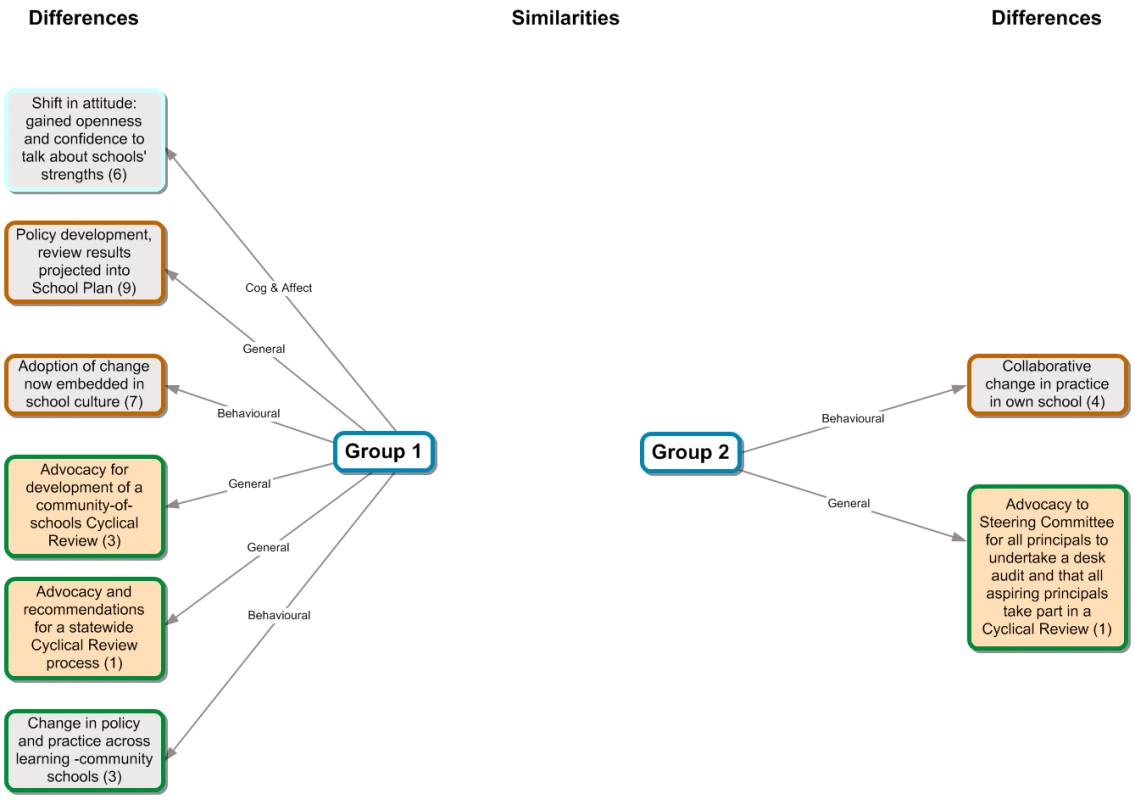
No unintended influences emerged for either group at this level. Three principals from Group 1, however, noted that because they had led a review team in one of their learning-community schools or because they had staff on a learning-community-school team, learning from the results had occurred right across the learning community and hence for this subgroup the behavioural influence of a change in policy and practice across schools occurred. As Principal F commented:

> *Having staff on our own team and on learning-community teams has facilitated change in practice—better sharing of ideas and better planning at learning-community level.*

Future directions from the reviews were also evident in the strategic plan for this learning community.

A summary of the long-term results-based influences on participating principals is presented in Figure 7.5 below.

Figure 7.5 Long-term Results-based Influences on Participating Principals



Process-based Influences on Principals

Principals as individuals

It was intended that as a result of leading and hosting a review, principals in Group 1 would not only have acquired new skills but be demonstrating them. It was further intended that their own attitudes and practice with regard to evaluation would be in line with the principles and procedures underlying the Cyclical Review process. These general, cognitive and affective, and behavioural influences did emerge for six of the principals, but to varying degrees.

For example, Principal E reported that over time she had come to realise the importance of the desk audit, had subsequently worked with a school development officer to understand the instrument better, and was now using it herself to assist her with her own school's self-evaluation. Principal F commented on the benefit of the experience with Cyclical Reviews in the school self-evaluation processes.

Principal G reported that she now understood the importance of a good review timetable, stating:

> *I didn't really understand the importance of a good timetable until the review got underway . . . the need for time in between to reflect and download data. It's a really important feature for the school and the team. The school gets better stronger data if the timetable works so there is the possibility that findings will be more explicit. The team gets a richer picture and time to reflect and discuss so there is more chance of professional learning and learning from the process. I didn't see all this until I had seen the reviews from both angles.*

Following the pilot reviews, this same principal led further Cyclical Reviews, observations of which confirmed that her involvement in the pilot had had a very positive influence on her leadership a year later, when the reviews were well led, the

findings were explicit, and the team clearly understood their role. This coheres with another principal's observation that:

> *The reviews build capacity for principals as team leaders'. I saw it with the Cyclical Review in my own school and experienced it as a leader. I don't mean only in evaluations but in the way principals work with teams at school.*
>
> (Principal A)

In summation, six principals confirmed that the reviews had an ongoing influence on them personally because they respected the process. They were prepared to learn new skills, change their attitudes and practices, and motivate staff to take part and trust the findings, because the process and outcome were inextricably linked and the process was sound.

Nevertheless, three principals did not report any lasting influence from their participation in the process. Although none claimed a detrimental effect, each remained silent on the matter.

Six individual comments from Group 2 principals suggested some ongoing influence from the review process a year or more after their involvement. With two exceptions the influences tended to pertain to individual principals rather than the group as a whole.

For Principal K the ongoing influence was general. He had introduced faculty reviews in his school and continued to reflect on and think more deeply about the process, determining that for him faculty reviews were both a developmental and accountability exercises, whereas the Cyclical Review was meant to be a developmental one only. If, however, he had concluded that both were good and both had their place, he had not committed his school to a Cyclical Review by the time of the final data collection.

Principal P, who intended to use the *Exemplary Practice Statements* and analyse his school but had not done so by the time the end-of-cycle data were collected, continued to report that he still intended to use the statements and analyse his own school. For this principal there appeared to be a motivational influence that had not translated into changed practice.

For a third principal the influence was cognitive and affective, as she believed that with time to reflect and the opportunity to be on another review team after the pilot, her attitude towards and understanding of reviews had changed. She commented that:

> *Participation brought home that if we want [a review] in our school we need to be specific, to focus. That means we need to really understand our school, where we are up to. I got this the first time but it really came home when I did a second one after the pilot and the principal was not as focused. You also need your staff on board before you do it. I sort if knew this but doing the second one brought it home. It's when that happens that you realise and then it impacts on what you do back in your own school and how you do it.*
>
> (Principal R)

Principal S reported that the quickness of the process reinforced the need for time to be made at a school level to communicate, come back, and discuss, and that this had resulted in a behavioural change in the way he now approached planning.

There was one influence, however, that was both cognitive and affective and behavioural and was reported by all of the above principals. They all asserted that the process cemented the understanding that data are critical. As Principal S said, '*this is the thing that still influences me and my practice and now my school almost two years on*'.

There were no unintended influences observed or reported for either group. What did emerge, however, in this long-term phase was an increasing difficulty to separate influences that occurred for principals in Group 1 as individuals and as school leaders.

As Principal D said, '*the process has become so embedded that it is difficult to see how it just affects me*'.

Principals as school leaders

It was intended that Group 1 principals would develop their skills in evaluation processes and that this would be reflected in the long term by their ongoing commitment to and leadership of rigorous data-based evaluation processes within their schools. As previously discussed one intended behavioural influence based on the process was therefore a demonstrated change in practice, showing the schools' increased capacities to conduct and report on management and curriculum evaluations as evidenced in these sections of the schools' annual reports.

As seven reviews took place in 2007 it was possible to analyse these schools' 2008 annual reports to determine whether the improvements—evident in their 2007 report, owing to the mandatory reporting of the Cyclical Review—continued, thus showing that the intended behavioural influence of a change of practice, indicating long-term increased evaluation capacity, had occurred. In three cases, there was clear evidence that a process similar to that of a Cyclical Review had become embedded into school practice.

The reports showed that one of the principals (Principal G) had used the identical process at a school level. She allowed eleven staff members to be trained and to participate in her own school's review, which was a whole-school review examining three areas that were not covered in the school's Cyclical Review. She reported on this in her school's annual report, writing:

> *In 2008 our school carried out evaluations of School Purpose, Performance, Professional Practice, Management, Organisation and Numeracy. These areas were combined with the findings from the Cyclical Review in 2007 in Teaching, Learning, and Curriculum*

to contribute to the 2009–11 school plan. Staff, students, and parents were interviewed to gather information.

Two other principals had adapted the process so that groups of staff within their schools were trained in the methodology and used it to evaluate and report on faculty and school operations. As one of these principals further commented:

> *The Cyclical Review model is far superior to other models, it builds capacity. We've used methods from it ever since. In fact there has been a long lasting positive impact on the way we do business—facts, data, listening to each other, openness, transparency.*
>
> (Principal A)

In two other cases the 2008 annual reports showed that some of the Cyclical Review principles were applied to one of the evaluations, with more analysis of collected data accompanying the stated future directions. One of these principals further commented:

> *The level of professional discussion continued amongst the staff. It has had a long term positive influence, a mixture of process and results. I use this to promote the school's directions and apply ideas to further evaluations.*
>
> (Principal I)

While acknowledging that these changes had occurred, principals were nevertheless cautious about attributing all the changes exclusively to their involvement in the Cyclical Reviews. Comments such as '*it's difficult to know to what the extent these changes are due to the review or to other factors; you know schools are complex places and there are lots of things happening and influencing our practices*' (Principal A) were common.

Disappointingly, however, the intended change in evaluation practice was not evident in all instances. In two cases there was no evidence of a change in the way these schools carried out or reported on their evaluations in their annual reports, instead reverting to descriptions of a strategy or program, as had been evident in the reports preceding the Cyclical Reviews. Furthermore, these principals did not report any influence as a result of the process in this phase of data collection.

It was intended that the Group 2 principals, acting as change agents, would continue to refer positively to Cyclical Review processes and, in fact, adopt of some of the evaluation processes or a full Cyclical Review in their own schools for their following year's mandatory evaluations, thus affecting a behavioural change in school practice.

In five cases, and to varying degrees, a behavioural influence emerged in this phase, although rather than being a discrete influence it appeared to be a combination of influences that culminated in diffusion: that is, the adoption of the Cyclical Review process or elements of it into their own schools. For example, Principal M reported that he had used his newly acquired knowledge and skills to get staff to do evaluations without the first resort being a survey. The annual report for the school in 2008 confirmed this to be the case, explaining that a team approach had been adopted and that the evaluation included the analysis of a wide a range of documents.

Two principals reported that although they had not used the full evaluation process, they and their staffs now used parts of it. One of them further commented that he had been motivated by the process and that as a result he and his staff:

> *Now look for data to back up observations and look to see where data are coming from—that is coming from more than the teachers. Our school evaluations are much more dependent on data and triangulation. I am much more aware of need for this and check that executive and staff can back up their plans and outcomes. General staff meetings and minutes, for example have agendas devoted to evaluation and data analysis.*
>
> (Principal S)

Again, the 2008 annual reports of both schools confirmed that a range of evaluation processes, similar to those of a Cyclical Review, were now in place. For example, in an evaluation of student learning:

> *Student skills were coded and surveys were conducted to determine whether students' skills had improved two terms after the strategy was implemented. The coding showed students continued to improve over time. The teachers appreciated the professional learning that occurred as a result.*
>
> (Principal S)

A fourth principal reported that he had been able to introduce internal faculty reviews led by his executive using similar processes to Cyclical Reviews, and his annual report also confirmed this to be the case. He commented that:

> *The faculty reviews use the same processes as a Cyclical Review although we do not use pocket PCs. However, we still collect, code, and triangulate all the data. We intend to rotate through each faculty. In 2009 it was science and in 2010 it will be English and the head teacher, science will lead. Then the English head teacher can lead the next one and so on.*
>
> (Principal K)

Finally, Principal Q, who, immediately following his participation in a review, had been motivated to gain the support of his staff and commit his school and staff to their own Cyclical Review, reported that staff had discussed the *Exemplary Practice Statements* and initiated their own discussion from the statements at meetings in preparation for their own review. He believed that this would change the way evaluation was carried out.

Disappointingly, influences based on the process did not occur in every instance and, in fact, some principals reported no influence or were silent on the matter.

One unintended influence to emerge, although only with four of the Group 1 principals, was a general one: the principals' deeper consideration of the widespread use of the *Exemplary Practice Statements*. While these principals reported that they continued to use the statements and thought that they were essential, especially in conjunction with the management plan and for ongoing executive development, they also were now convinced that staff, especially early-career teachers, would understand

better the professional teaching standards rather than the *Exemplary Practice Statements*. No unintended influences were observed or reported for Group 2 principals.

Principals as system leaders

For Group 1 principals one general and one cognitive and affective intended influence emerged. The general influence was advocacy beyond the pilot group for Cyclical Reviews and the cognitive and affective influence was a call for modification of some of the processes.

It was intended that within one to two years of the pilot, Cyclical Reviews would be embedded throughout Western Sydney Region schools and that this would be due, in part, to participating principals spreading the word about their own positive experiences and persuading others to become involved. At the time of the final data collection this had not happened to the extent that was intended. Five of the principals, most notably those who were also on the Steering Committee, took opportunities to speak about the process and to convince others to become involved. For example, Principal F reported that she had recommended Cyclical Reviews to her colleagues at principals' association meetings, not only because of the benefit to the school but also in terms of principals' professional learning. Three others reported that they had mentioned their involvement to their colleagues on occasions, but had not specifically tried to persuade others to become involved. So while principals were influenced to varying degrees to act as advocates, a change in practice at a regional level did not eventuate, at least not to the extent that was intended.

During the conduct of the reviews principals had commented on a range of factors that either helped or hindered the review process (as discussed in Chapter 6), but it was not until this long-term phase that principals decided to examine the data collected from the reviews, conduct some research into other systems' practices, and from this modify

some of the processes, most notably the *Exemplary Practice Statements* and the desk audit. One principal, who had taken a very active role in the initial development of the statements, voiced the group's view when he commented that he was:

> *Open to redevelop the new Exemplary Practice Statements as a result of feedback from reviews, especially the results that had been reported. There is a need to modify the statements to get more explicit findings. There is a need now to take into account the groups' learning as well as learning that has occurred in other systems that might be of value.*
>
> (Principal A)

For Group 2 principals the one intended general influence, that of advocacy to other principal colleagues to host a Cyclical Review, did not really emerge. Although five principals reported that they told other principals that the experience as a team member was beneficial, only Principal Q reported and was also observed advocating that his colleagues consider a Cyclical Review in their own schools. The one behavioural influence to emerge, however, that applied to eight of the principals in Group 2 was that they sought opportunities for members of their staffs to take part in other review processes that were available across and beyond the region, because they saw the value in these as professional-learning opportunities.

In accord with their Group 1 colleagues, a cognitive and affective influence emerged for four principals in Group 2. They also called for modifications to some of the processes. For example:

> *The more I think about the Exemplary Practice Statements the more they need revising. Not because they are not right but because we need to make them more manageable. What do other states do? What happens overseas? People need to go and find out and bring back to a meeting so that we can decide where to go from here.*
>
> (Principal S)

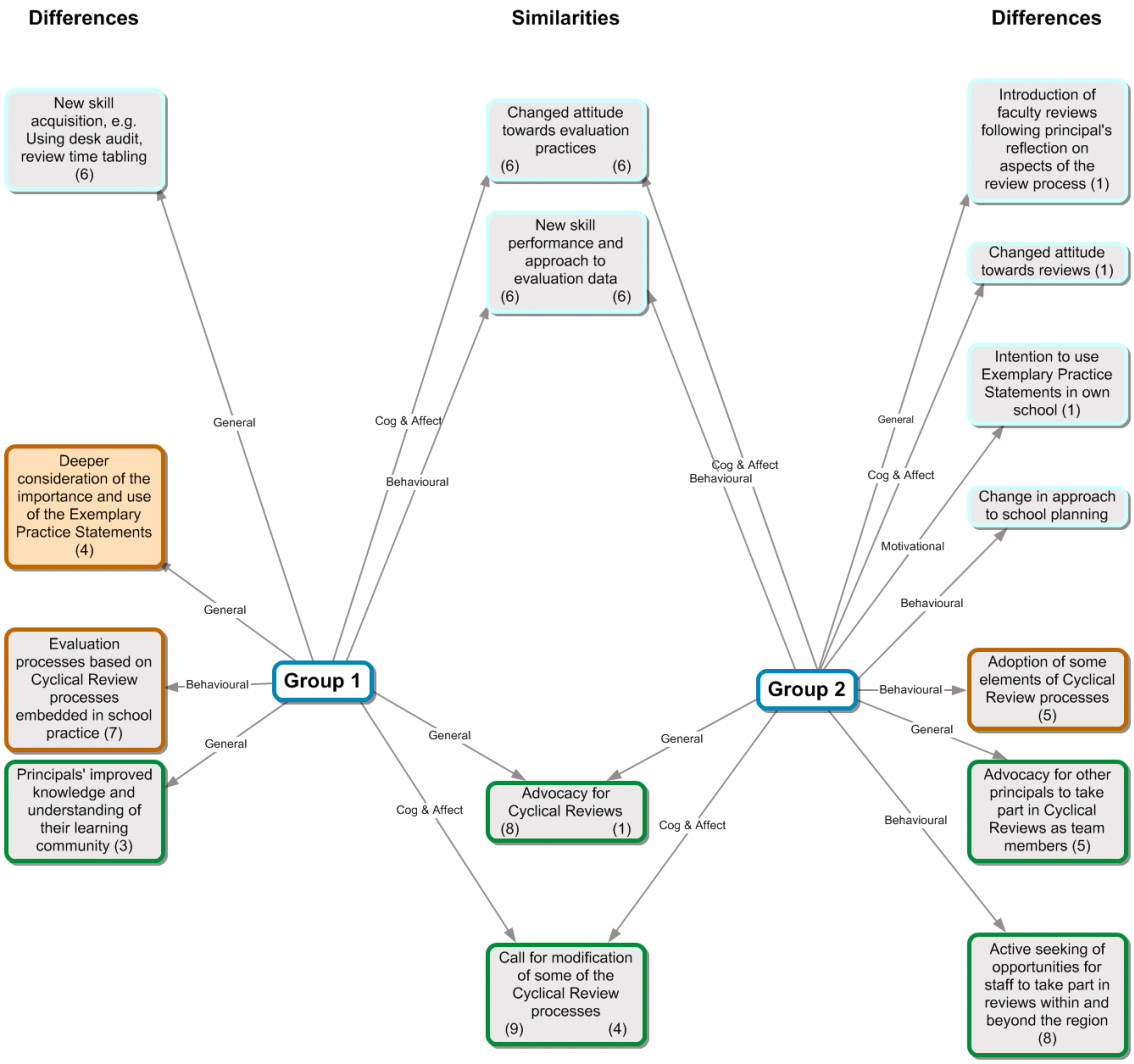
For Group 1, one general influence did occur that was unintended and was manifested in two different ways, although it must be said that it did not apply to the group as a

whole. First, three of the principals who belonged to the same learning community requested that the Steering Committee consider developing a modification of the process whereby a Cyclical Review could be undertaken across a community of schools. Second, one of the principals wrote a position paper on Cyclical Reviews for the state body of the Secondary Principals' Council, recommending the process for statewide implementation. His recommendations and those of others on the Steering Committee were being considered by a Departmental committee at the time of writing to shape a cyclical review process for the state.

Although no unintended influence emerged that applied to Group 2 as a whole, one principal, who by this time had taken part in a second Cyclical Review beyond the pilot, was influenced to the extent that he recommended two initiatives to the regional leadership group. The first was that the desk audit was so important that all principals should undertake it by tying it into the Principal Accountability and Review Schedule (PARS). The second was that participation in a Cyclical Review should become mandatory for anyone aspiring to be a principal or an executive, because it was an authentic exercise.

A summary of the long-term process-based influences on participating principals is shown in Figure 7.6 overleaf.

Figure 7.6 Long-term Process-based Influences on Participating Principals



IMPLICATIONS

Results-based Influences

The findings in Figures 7.1, 7.3, and 7.5 above were examined for results-based patterns of influence. The first pattern to emerge was the concentration of results-based influences over time for Group 1 principals compared with Group 2 principals.

Within this pattern, there are four strands. The first strand concerned leadership actions. In Figure 7.1 the immediate influences of the review results on principals as individuals were the adoption of the *Exemplary Practice Statements*, improved knowledge of their staffs and schools, and strengthened convictions about directions. By the end-of-cycle phase, shown in Figure 7.3, these influences had become explicit in leadership actions and in the acceptance of system requirements regarding school accountability, school planning for the following year, and evidence-based practice. In Figure 7.5 these developments were continuing at a school-leadership level as policy development and the enculturation of the changes developed from review results.

The second strand, also noted in Figures 7.1, 7.3, and 7.5, was about the principals publicly accepting the legitimacy of the results of the reviews. In Figure 7.3 and 7.5 this moral legitimacy underpinned ongoing school leadership activities.

The third strand noted in Figure 7.1 was the emergence of system leadership by the principals who belonged to the same learning community. This re-emerged in Figure 7.5 as changes in policy and practices they coordinated.

A fourth strand at the end-of-cycle phase emerged with three distinctive features: the prioritisation of findings for attention; an intensification of focus against standards; and findings achieving high saliency in professional discourse. One long-term legacy noted in Figure 7.5 was the shift in attitude to talking openly and confidently about schools' successes. Another was the embedding of changes as the 'new normal' in schools' cultures.

All of the above strands were intended in the design of the Cyclical Review process. One unintended results-based influence occurred when two principals used the reviews and the results to justify decisions they had already made, as shown in Figures

7.1 and 7.3. Another unintended outcome in the long term was a demand by principals as system leaders to create a Cyclical Review process for school-learning communities.

The analysis to this point reinforces Kirkhart's (2000) position that an influence that emerges early can lose saliency and then re-emerge in the long term.

The second pattern to emerge was the similarities between Group 1 and Group 2 principals with regards to results-based influences over time for Group 1 principals compared with Group 2 principals. Although unintended as an immediate results-based influence, it is evident in Figure 7.1 that some review-team principals in both groups improved their knowledge and understandings about their own staffs and schools. In Figure 7.3, as anticipated, both groups of principals used the results from the reviews they had conducted to reflect on achievements and expectations in their own schools.

It was interesting that six principals in Group 2 and one principal in Group 1 reported that the impact of the review report was immediately dependent on its specificity, as shown in Figure 7.1. This unintended influence was probably due to Group 1 principals being closely involved in the design of the reporting template.

A third pattern to emerge and which was an unanticipated immediate influence was a surge of advocacy in both groups for recognition and dissemination of exemplary practices. This largely symbolic action was not evident at the end-of-cycle and the only call for advocacy in the long-term was by one principal from Group 1 (who advocated for statewide Cyclical Reviews) and another principal from Group 2 (who called for all principals to undertake a desk audit and aspiring principals to be on a review team).

A fourth pattern, common to both groups, again unintended, that emerged at the end-of-cycle, was when principals applied findings from other schools' reviews and

applied it to their own schools. What had been intended and, in fact, emerged was that principals on review teams from both groups (all nine Group 1 principals and five Group 2 principals) became change agents in their own schools. Four of these five Group 5 principals went on to provide collaborative change practices in their schools in the long-term. Similarly, seven of the nine Group 1 principals claimed to have embedded collaborative change agency in their own practice in the long term (see Figure 7.5).

Process-based Influences

The most immediate patterns to emerge from Figures 7.2, 7.4, and 7.6 were:

1. the substantial similarities in immediate influences on both groups of principals
2. the marked bifurcation of influences at the end-of-cycle and this bifurcation being sustained in the long term
3. the very large number and diverse nature of influences on Group 2 principals at all three phases compared with the more focused and analytic approach adopted by the Group 1 principals
4. the very low number of unintended influences to emerge for both groups
5. the gradual emergence of system leadership influences.

Taken together, these patterns appear to illustrate the marked difference in the degree of learning about school evaluation achieved by the two groups of principals. Group 2 principals continued to use a wide range of personal theories about evaluation and tended to use processes on an ad hoc basis. Group 1 principals became far more analytical as a result of their early ownership of and their subsequent engagement in coherent review processes.

To illustrate the four points above in greater detail, Figure 7.2 shows that both groups of principals immediately displayed greater attention to evaluation practices, increased

knowledge and skills about evaluation, attitudinal change towards evaluation design, motivation to learn, and demonstrated newly acquired skills.

At the end-of cycle, as shown in Figure 7.4, the Group 2 principals were still displaying very similar process-based influences compared with the immediate phase influences. The exceptions were that by the end-of-cycle there was no further evidence of their attention to evaluation design and in two cases they had applied the *Exemplary Practice Statements* in their own schools. Seven of the nine Group 2 principals were using evidence increasingly in their own schools, which may have been a legacy of the newly learnt skills, as highlighted in Figure 7.2. Further, four principals from Group 2 reported acting as change agents, eight were advocating Cyclical Reviews, and one was motivated to commit to a Cyclical Review.

In sharp contrast, by the end-of-cycle the Group 1 principals had moved on into embedding cultural change through the more logical and systematic processes promoted by Cyclical Reviews. At the same time Group 1 principals were reporting changes of attitudes towards staff members on the review teams and more open and transparent practices.

This bifurcation in the long term continued with a few exceptions. The exceptions included continuing changes in attitudes towards evaluation practices and the acquisition of new skills and performance and approaches to do with evaluation data.

The unintended influences at the immediate phase applied only to Group 1 principals. The first was the perceived influence of the Regional Director. The second was anxiety at having to perform newly acquired skills. The only end-of-cycle influence that applied to a very small number of principals in both groups was the ongoing personal use of the pocket PC. The only unintended long-term influence, again for Group 1 principals, was deeper consideration of the *Exemplary Practice Statements*.

Regarding system leadership, both groups advocated Cyclical Review processes beyond the school. This endured through the end-of-cycle phase, but in the long term had split into advocacy for two separate influences. The Group 1 principals tended to advocate for Cyclical Reviews, while the Group 2 principals advocated for other principals to take part in Cyclical Reviews as team members. In each case they were advocating for others to have the experience that they had had. On the other hand, both groups were calling for modification of some of the Cyclical Review processes and most Group 2 principals were seeking opportunities for staff to take part in reviews within and beyond their region. This can be taken as indicating further interest in directing their own and others' professional development. Finally, unique to Group 1, was a minority reporting that they had improved knowledge and understanding of the schools in their learning community that had also taken part in the reviews.

Interaction between Results-based and Process-based Influences

By comparing the immediate results-based influences with process-based influences in Figures 7.1 and 7.2 it can be seen that there were substantial similarities in the nature of learning for principals in both groups. On the other hand, the start-up requirements of a Cyclical Review process required Group 1 principals to develop an appreciation of the *Exemplary Practice Statements*, systematically analyse their staffs and their schools, clarify their schools' directions, and publicly commit to implementing findings. In contrast the Group 2 principals only had to commit to participating in a school-review team. The other great difference between the two groups was that Group 1 principals helped design the operational details for the Cyclical Review process while the Group 2 principals merely participated in the review teams. This meant that Group 1 principals were actively engaged in designing their own learning and that this learning was further deepened by participation in the review team and gaining feedback in a learning experiment in their own schools. Put simply, Group 1 principals used double-loop learning to design and consolidate their learning about Cyclical Reviews, while

Group 2 principals participated in another school's review without having the challenges and advantages of gaining empirical feedback about the effectiveness of their understandings of school evaluation.

There is a methodological limitation that must be highlighted at this point. The analysis of results-based influences was not fine-grained enough to distinguish between the effects of being a host principal or a combination of being a host principal and a team leader. On the other hand it seems to be implausible that the effects in Group 1 were due solely to being a host principal. Further research will therefore be required to isolate the effects of merely hosting a Cyclical Review from the effects of both assisting with the design and hosting a review and reflecting critically on the outcomes. What appears to be the case is that process-based influences prefaced results-based influences in the immediate term.

By the end-of-cycle phase, as evidenced in Figures 7.3 and 7.4, the bifurcation in process-based influences had become particularly marked, primarily due to the rigour of the Cyclical Review processes and the learning about evaluation that they stimulated.

The long-term influences, summarised in Figures 7.5 and 7.6, shows that Group 1 principals had embedded Cyclical Review evaluation culture and practices into their schools. In contrast Group 2 principals were still grappling with the need to develop a comprehensive model of evaluation, with adoption characterised by the ad hoc or incoherent use of methods. Again process-based influences prefaced results-based influences. Yet again it was not possible to discriminate between the effects of being a host principal and a combination of a host principal and a team leader.

What can be asserted on the basis of this research is that early and deep learning about evaluation processes are essential to gaining the full effects of Cyclical Reviews.

Group 1 principals engaging in deep learning created both fresh structures in their schools and fresh outcomes, illustrating Giddens's (1984) theory about the duality of structure: it is both process and outcome.

Implications for Theory

The above discussion about process-based and results-based sources of influence has implications for components of the theory discussed in Chapter 3. These links will now be briefly discussed before wider provisional conclusions are drawn in the final chapter.

Kirkhart (2000:7) theorised that the impact of evaluation is maximised when it shapes, affects, supports, and changes people and systems simultaneously and she therefore stressed the importance of influence in all of these realms. She went on to propose an integrated theory of influence that addressed the dimensions of source, intention, and timeframe. Regarding source, she recommended attention to both process and results. Regarding intention, she recommended attending to both intended and unintended influences. Regarding timeframe, she recommended distinguishing between immediate, end-of-cycle, and long-term phases. All of this theory has been substantiated by the research outcomes reported here.

It was explained in Chapter 3 that Henry and Mark had proposed improvements to Kirkhart's thinking. They began by modelling alternative mechanisms (Mark & Henry, 2004) that may mediate evaluation influence. These mechanisms proved a useful and finer-grained checklist when analysing the findings reported in this study. They helped in Figures 7.1–7.6 to classify the types of influence as being general in nature, cognitive and affective, motivational, or behavioural. While these types were helpful in understanding interactions, the schematic theory of evaluation influence proposed by Mark and Henry was dependent on an input, throughput, and output framework that

did not cohere with the interactive nature of the PAR used in the Cyclical Reviews reported in this study.

The findings on process and results-based influence strongly cohered with Alkin's (1985) original portrayal of how context and human factors interacted with evaluation processes. This study suggested a correlation between these three factors and that the interplay between these three factors later influenced both process and results.

The literature also stressed that Cyclical Reviews were a form of participatory evaluation with all of the features of PAR. Cousins and Whitmore's (2007:93) model of dimensions of form in collaborative inquiry used three dimensions: control of evaluation processes; stakeholder selection for participation; and depth of participation. The discussion above of evaluation influences on Group 1 principals indicated that they had a high degree of control of Cyclical Review evaluation processes; that is, principals as researchers as well as principals as practitioners. This did not apply to Group 2 principals. With regard to stakeholder selection for participation, the review teams deliberately engaged a range of legitimate groups, in one case including a parent, and this could be expanded in future. With regard to depth of participation the Group 1 principals were deep participants, whereas Group 2 principals only played a consultation role. In Cousins and Whitmore's terms the evidence above is that the patterns of influence from Cyclical Reviews were a consequence of the combination of principal-as-researcher controlled process, wider stakeholder selection for participation, and deep participation. This means that the parallels supposed between P-PE and Cyclical Reviews were not verified and that a more sophisticated explanation was needed with regard to Cyclical Reviews because they had a dual purpose: that of practical problem solving in each school to inform decision-making; and the development of evaluation capacity.

ECB was described as organisational learning processes intended to make quality evaluation routine (Baizerman et al., 2002b). Argyris and Schön (1978) argued that

evaluation had to help individuals, groups, and whole organisations move from reacting to differences between expected and actual outcomes (single-loop learning) to examining the deep values, assumptions, and policies that lead to actual outcomes and to renegotiate them (double-loop learning). Organisational development theory (King & Volkov, 2005) focuses on the information systems used to learn about changes in the environment, improving evaluation to provide feedback for deep learning and boosting trust between the staff of a school.

The above discussion, summarised by Figures 7.1–7.6, offers many examples of how the Cyclical Review processes and outcomes deliver ECB. The sole exceptions concern the misuse of Cyclical Reviews by two Group 1 principals, who used them to justify decisions already made, a form of pseudo-evaluation discussed in Chapter 3. On the other hand, the ability of the Cyclical Review process to deliver both practical solutions and ECB illustrates Giddens's (1984) theory about the duality of structures; they are both process and outcome. In this case the Cyclical Review processes delivered new structures in schools and new professional leadership capacities in evaluation.

Implications for Policy

In this section the potential challenges of embedding a policy of Cyclical Reviews in the New South Wales government school system is discussed.

The above discussion concerning patterns of influence shows that principals must engage in process design to boost ownership and deep learning in order to gain the full effects of participation in Cyclical Reviews in the medium and long term. It also indicates that this engagement exposes them to the rigour of action research that provides evidence-based feedback customised by principal and school. These attributes will need to be retained if the same influences are to be a feature of a system-wide Cyclical Review policy.

One of the reasons given by principals for the success of the Cyclical Reviews was that they were not mandatory. This poses a policy challenge. Making them mandatory would violate conditions considered essential by participating principals. Leaving them voluntary may allow some principals to evade discharging their public accountabilities under current policies. In Chapter 2 it was explained that New South Wales currently has the *School Development Policy* with three types of non-mandatory school reviews: management reviews, program reviews, and educational support team visits, with Cyclical Reviews designed to adhere to the principles of the third of these.

One possible policy solution is to implement Cyclical Reviews, while retaining reviews by exception based on the *School Development Policy*. This would mean aiming at embedding Cyclical Reviews as the norm through a structured training and accreditation program aimed at ECB for principals as part of the more recent *Local Schools, Local Decisions* implementation.

There is the possibility that locally agreed sets of standards may be replaced by national standards and declared as fixed policy. This would run against the need to have principals design frameworks of standards during the start-up of Cyclical Reviews in their defined area, albeit refined versions of prior standards. This philosophical exercise is the first step of ECB in defining purposes and criteria for evaluation. It is the first phase of deep learning and an essential precursor to double-loop learning.

Ideally, this also suggests that the management of Cyclical Reviews should be in the hands of Group 1 principals. A practical substitute at system level would be to establish a training and accreditation program that imparted Group 1 experiences and deep learning.

Implications for Practice

The patterns of influence suggest a number of operational improvements for Cyclical Reviews:

- the further development of the desk audit, survey instruments, interview schedules, reporting templates, etc.
- the replacement of all aspects of PPODS and instrumentation to embrace new and contemporary technologies so that data can be collected, sorted, and captured more quickly, accurately, and effectively
- identifying competencies and training that lead to the accreditation of the four levels of participation: coach, team leader, host principal, and team member
- the refreshment of criteria and processes for selecting participation in review teams.

SUMMARY

This chapter presented answers to the third and fourth research questions:

- How does participation in Cyclical Reviews in Western Sydney Region influence participating principals?
- To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?

Participation influenced principals according to their roles. The different degrees of deep learning, the coherence of evidence collected and applied, and the pace at which cultural change was embedded were found to be related to the nature of their participation in evaluation processes and the subsequent acquisition of new personal theories and practices.

Similarly ECB was far more evident in Group 1 principals over time than in Group 2, largely as a consequence of the extent to which they engaged in deep learning about evaluation and double-loop learning activities.

In the next and final chapter, the tentative answers to the first research question, 'is there a theoretical model that can be designed to map evaluation influence?' will be developed, along with the tentative conclusions of this study, a summary of the implications for theory, policy, and practice, and suggestions further research.

CHAPTER 8
CONCLUSIONS AND RECOMMENDATIONS

INTRODUCTION

Tentative conclusions based on the findings of this study are presented in this final chapter, along with recommendations for practice, policy, theory, and further research. To reiterate briefly, the purpose of the study was to determine the influence the Cyclical Review process had on participating school principals. It was expected that insights would be gained into the use made over time of both the evaluation processes and evaluation results by the participating principals. The aim was to gain impressions from the principals about the factors, if any, that had influenced them to use the evaluation processes and results both in the course of the review and for at least twelve months afterwards. The practical intention was to use this knowledge to modify and strengthen the Cyclical Review program in Western Sydney Region and, in so doing, contribute to the statewide implementation of that element of the New South Wales *School Development and Accountability Framework*. Finally, it was expected to add to current knowledge concerning Cyclical Reviews both as theory and policy.

A qualitative approach used case-studies to collect data from multiple sources. Two groups of principals were compared: the first was involved in the design of Cyclical Review criteria and processes, hosting reviews, and leading reviews; the second served only as team members on others' reviews.

The approach also required the development of a provisional theoretical model to classify the data to address the research questions and interpret the findings in terms of the relevant literature. A process developed specifically for the Cyclical Reviews, PPODS, was used to collect, enter, and sort the data.

The literature referred to focused on educational evaluation; its functions and forms, evaluation utilisation, use, and influence; factors affecting evaluation; participatory evaluation; and ECB.

Two limitations must be highlighted again at this point. The first, referred to in Chapter 6, was one of scope. That key values act as threshold conditions for change was an important finding to emerge. Nevertheless, while the body of literature pertaining to values and leadership is important, it was acknowledged that this was beyond the scope of the study and would warrant further research.

The second, referred to in Chapter 7, was methodological. The analysis of results-based influences for Group 1 principals was not fine-grained enough to distinguish between the effects of the roles they played: those of being a host principal or a combination of being a host principal and a team leader. On the other hand, the marked difference in the results between Group 1 and Group 2 principals provided support for the tentative conclusion that the effects in Group 1 were not due solely to being a host principal. Further research will therefore be required to isolate the effects of merely hosting a Cyclical Review from both assisting with the design and hosting a review and reflecting critically on the outcomes. What can be said with some conviction is that process-based influences prefaced results-based influences in the immediate term. In other words, the study suggests a causal link between process-based and results-based influences that could be explored through further research.

With these caveats in mind, this final chapter begins with the tentative findings to answer the first research question, 'is there a theoretical model that can be designed to map evaluation influence?' Subsequent sections will present findings on the other research questions, offer tentative conclusions, and a summary of the implications for theory, policy, and practice. The chapter will end with suggestions for further research.

THE RESEARCH QUESTIONS

The research questions arose from four major recurring themes introduced and contextualised in Chapters 1 and 2 and then detailed by an examination of the literature in Chapter 3. These themes were: (1) a model to record and analyse the data in order to determine evaluation influence; (2) knowledge, prior experiences, and factors that influence principals who participate in evaluation processes; (3) the distinction between evaluation use and evaluation influence; and (4) Cyclical Reviews as ECB.

The discussion of the theoretical models of evaluation influence in the literature and the development of a new model that would be needed to undertake this study led directly to the first research question:

1. Is there a theoretical model that can be designed to map evaluation influence?

Second, the majority of the scholarly work reviewed, both in the practical context of the study and in the research literature, identified the factors that contributed to effective evaluation for accountability or developmental purposes or a combination of both. Given the third purpose of the study—to modify and strengthen the Cyclical Review process—, the initial factors that trigger any subsequent influence needed to be determined. This led to the second research question:

2. What factors, prior experiences, and understandings contribute to the influence that the involvement in Cyclical Reviews in Western Sydney Region has had on the participating principals?

Third, as the distinctions between evaluation use and evaluation influence were drawn from the literature, it became evident that if the first and second broad purposes of the study were to be pursued, then the latter and more comprehensive concept of

evaluation influence was more appropriate for this study; it subsumed the former and less useful concept of evaluation use. This gave rise to the third research question:

3. How does participation in Cyclical Reviews in Western Sydney Region influence participating principals?

Fourth and finally, the literature review analysed the theoretical grounds, definitions, and processes of ECB and explored the links between self-evaluation, OD, and organisational learning. Given the intentions of the Western Sydney Regional Director in commissioning the development of the Cyclical Review process, and the congruence between his intentions and the literature pertaining to ECB, the final research question was:

4. To what extent are the outcomes of ECB demonstrated by the principals who participated in the Cyclical Reviews?

FINDINGS

Is there a Theoretical Model that can be Designed to Map Evaluation Influence?

In order to mount this study a provisional model of evaluation influence was developed from the range of relevant literature including, in particular, the models proposed by Kirkhart (2000) and Mark and Henry (2004). This tentative model classified the factors that could trigger influence as context, human, and evaluation, in line with Alkin's (1985) categories. It also conceptualised influence on four dimensions: source (process and results); intention (intended and unintended); principals (as individuals, as school leaders, and as system leaders); and time (immediate, end-of-cycle, and long term). Together these elements initially appeared to explain the degree of influence in two areas: the application of review findings and ECB.

While the theory and model derived from it proved effective for data-gathering purposes in this study, a number of limitations were found with regard to the analysis of patterns of influence, one of the key factors identified as a core element from the literature. First, the model as depicted did not adequately represent how learning about evaluation influences differed between the two groups of principals involved in the study. Second, it did not represent the catalytic conditions that appeared to make a real difference to the degree of both results and process-based influence. While numerous factors and sub-groupings of factors were present, a set of catalytic conditions *per se* was not identified in the literature. Third, it did not recognise the role that values and assumptions were found to play nor the dynamics of structuration, whereby evaluation itself was improved and evaluation capacity was built. Fourth, while the provisional model enabled an episodic mapping and analysis of the micro-dynamics of influence, a more macro-model was required to map and explain the broader causal story of influence found in the data about Cyclical Reviews. These implications will be discussed in more detail throughout this chapter. Figures 8.1–8.3 that accompany this discussion will illustrate the development of the theoretical conclusion of this study: a proposed model of influence.

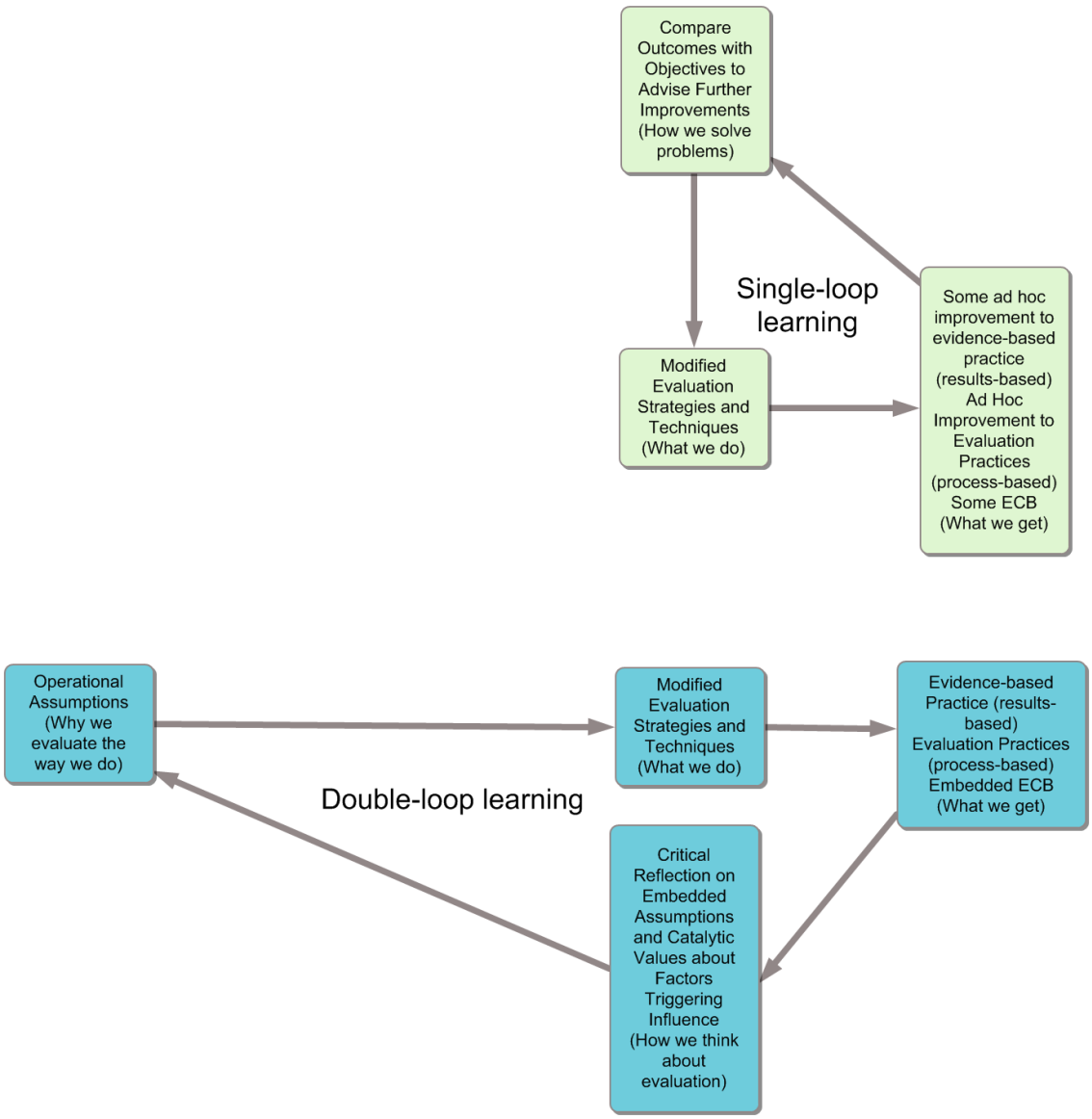
The basic schism in the findings presented in Chapters 6 and 7 was the difference between the general experience of influence experienced by Group 1 and Group 2 principals. Group 2 principals, who only advised as team members in others' schools, were primarily concerned with problem-solving when outcomes fell short of objectives. They modified their evaluations and strategies and techniques in order to make relatively ad hoc improvements to their own evaluation practices and capacity building. This was consistent with the limits of single-loop learning as theorised by Argyris and Schön (1978), Imants (2003), Senge (2010), and Senge et al. (2000; 2010) and discussed in Chapter 3.

In sharp contrast, Group 1 principals were strongly influenced by having to understand, apply, and improve criteria used in the evaluation of their own schools,

design evaluation strategies and techniques, and later reflect critically on the quality of evaluation practices and capacities in order to improve the values and assumptions influencing their own practices. These activities illustrated comprehensively the components of double-loop learning.

The basic structure of these two patterns of influence is portrayed in Figure 8.1 overleaf. The bottom half of the model (shaded blue) depicts the broad influence experienced by the Group 1 principals as double-loop learning. The top half (shaded green) depicts the broad influence experienced by the Group 2 principals as single-loop learning.

Figure 8.1 Basic Structure of Patterns of Influence



The panels in Figure 8.1 represent distinct sets of activities that together comprise a learning process, the arrows indicating the sequence. The basic model will be elaborated in Figures 8.2–8.3 below in the light of the answers to the other research questions and recommendations for theory.

What Factors, Prior Experiences, and Understandings Contribute to the Influence that the Involvement in Cyclical Reviews in Western Sydney Region has had on the Participating Principals?

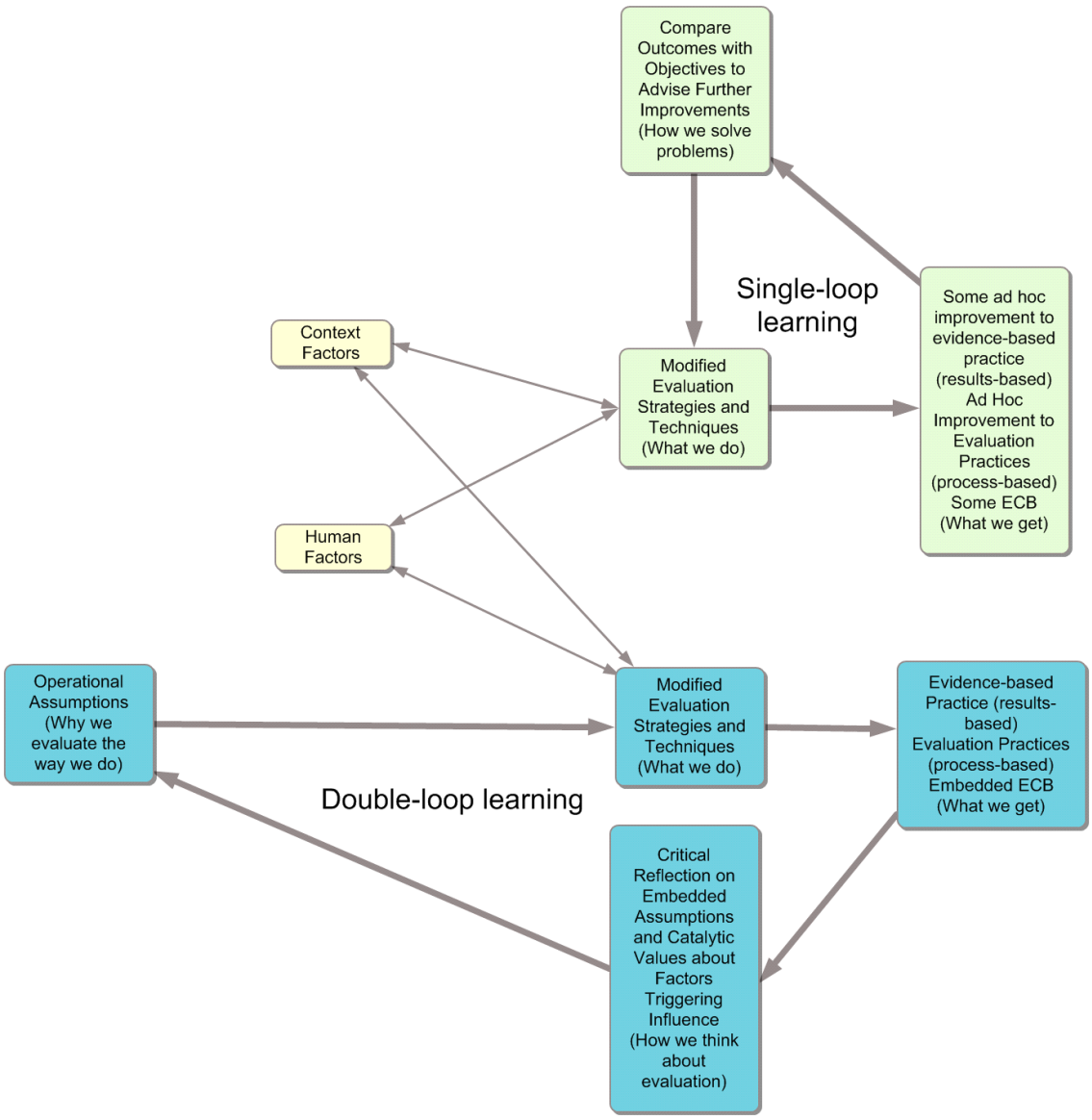
The factors, prior experiences, and understandings that contributed to influence on participating principals were clarified above in Chapter 6. They were grouped into three categories—context, human, and evaluation factors—in line with existing theories from the literature presented in Chapter 3. The context factors found to trigger influence included school culture, training, duration and timing, team size and composition, and review focus. The human factors found to trigger influence included motivations, and the interaction between the leadership of the review teams, the leadership by host principals, and human relationships. The evaluation factors found to trigger influence included structures and resources, the data-collection processes, and PPODS.

From the analysis and synthesis of findings presented in Chapter 6 it was found that the evaluation methodology used in Cyclical Reviews was primarily independent as a process but differentially affected by context and human factors in each school setting. The major variances to standardised review processes were the leadership services provided to the review team, the host principal, and the partnership between the review leader and the coach (whose role was discussed in Chapter 2). The conclusion drawn from this finding was that the Cyclical Review evaluation process varied substantially by school contextual factors and, in particular, by the interaction of the human factors associated with leadership. While this conclusion is itself an interesting aspect of the relationship between the three factors that trigger influence, it, in turn, now leads to two further conclusions.

First, it can be theorised that the evaluation factors incrementally influence the modification of evaluation strategies and techniques used in each setting, and can therefore be considered to be intrinsic to the activities represented by the panels

labelled *Modified Evaluation Strategies and Techniques (What we do)* presented in Figure 8.2 overleaf. Second, this means that context and human factors (shown in yellow in Figure 8.2) interact with evaluation factors and do so over time as participants make improvements to their evaluation practice. Their interactive relationship is indicated by the double-headed arrows. This is consistent with prior theory about single and double-loop learning as proposed by Argyris and Schön (1978). The key aspects of single and double-loop learning theory are that, through an entity's double-loop-learning attempts to reach a goal, the entity is able to modify the goal in the light of experience or even reject the goal. Through an entity's single-loop-learning attempts, however, the entity makes repeated attempts at the same problem, with no variation of method and without ever questioning the goal. It is interesting that more recent practitioners and theorists, such as Degenhardt and Duignan (2010), Hargreaves (2013), Degenhardt and McCulla (2013), and Stoll (2013), take for granted and regularly apply the ideas of double-loop learning.

Figure 8.2 Contribution of Context, Human, and Evaluation Factors



How does Participation in Cyclical Reviews in Western Sydney Region Influence Participating Principals?

It was concluded at the end of Chapter 7 that participation in Cyclical Reviews influenced principals according to the role they played, in line with participatory evaluation and its links with ECB, much as proposed by Owen (2006:220–221) and as discussed in Chapter 3. The different degrees of deep learning, the coherence of

evidence collected and applied, and the pace at which cultural change was embedded were found to be related to the nature of the principals' participation in evaluation processes and the subsequent acquisition of new personal theories and practices.

These differences are shown explicitly in Figure 8.2 as the differences between single-loop and double-loop learning. One example consistent with single-loop learning discussed in the literature occurs when a participant connects a strategy for action with a result and makes corrections. If during single-loop learning an action taken yields results that are different from what was expected, the participant will observe the results, automatically take in feedback, and try a different approach. This trial-and-error approach matches closely the learning activities that the Group 2 principals tended to use, as mapped in detail in Chapter 7.

In contrast, and consistent with double-loop learning discussed in the literature, one example of a participant's action would be to go beyond this basic approach and re-evaluate the deeper governing variables that appear to explain why participants behave in the ways they do. Re-evaluating and reframing goals, values, and beliefs are more complex ways of processing information and involve more sophisticated ways of engaging with an experience and look at consequences from a wider or more philosophical perspective. This again matches closely the sorts of learning activities and capacity-building strategies that the Group 1 principals engaged in, as discussed in detail in Chapter 7. This important conclusion coheres with the exploratory works of Senge, Scharmer, Jaworski, and Flowers (2008), Scharmer (2009), and Scharmer and Kaufer (2013), who have proposed 'U learning' (radically reconceptualising and reconstructing organisations) as critical in organisational learning.

To what Extent are the Outcomes of ECB Demonstrated by the Principals who Participated in the Cyclical Reviews?

It was further concluded at the end of Chapter 7 that the outcomes of ECB were far more evident in Group 1 principals over time than in Group 2, largely as a consequence of the extent to which they engaged in deep learning about evaluation and other double-loop learning activities.

As discussed and mapped in detail throughout Chapter 7, both groups of principals demonstrated ECB to varying extents. At the outset there were many similarities in the nature of learning about evaluation for principals in both groups; most notably their skills and understandings in evaluation were enhanced, and they began to demonstrate an improved ability to use evaluation data and processes. Over time, however, there was a marked difference in the two groups in the extent to which ECB outcomes were demonstrated.

By the end-of-cycle phase the bifurcation in process-based influences, and hence the extent to which principals had developed and demonstrated evaluation capacity, had become particularly marked. These differences were primarily due to the rigour of the Cyclical Review processes and the learning about evaluation that they stimulated. Group 1 principals, in general, had developed to a far greater extent the skills to plan, design, collect, and analyse data and to interpret and report results. The long-term influences showed that Group 1 principals demonstrated the outcomes of ECB identified in the literature; they had embedded Cyclical Review evaluation culture and practices into their schools. In Stufflebeam's (2002) terms, they had institutionalised reforms. In Giddens's (1984) terms, they had learnt to manage structuration. In contrast, Group 2 principals were still grappling with the need to develop a comprehensive model of evaluation with adoption characterised by the ad hoc or incoherent use of methods. This implies that Group 1 principals or other principals who gain similar experiences and skills could become coaches in future school-based

cyclical-review processes, offer additional coaching, or provide other leadership roles in similar contexts in the future.

Put another way, Group 1 principals used double-loop learning to design and consolidate their learning about Cyclical Reviews, while Group 2 principals participated in another school's review without having the challenges and advantages of gaining empirical feedback about the effectiveness of their own understandings of school evaluation. As a result the early and deep learning about evaluation processes experienced by the Group 1 principals led over time to their demonstrating ECB outcomes and gaining the full effects of Cyclical Reviews to a considerable extent. Further, Group 1 principals engaging in deep learning created both fresh structures in their schools and fresh outcomes, including those of new professional leadership capacities in evaluation, and thereby illustrated another aspect of Giddens's (1984) structuration theory about the duality of structure: it is both process and outcome.

The markedly different extent to which participation by the two different groups resulted in ECB is also illustrated in Figure 8.2. The implications of these findings follow in the next section, where they inform the development of recommendations.

RECOMMENDATIONS

Recommendations for Theory

In Chapter 6 it was concluded that significant advances in the tentative theory for this study were warranted by incorporating factors triggering influence into a broader model of how principals learn about evaluation and ECB. In this study, context and human factors interacted with evaluation factors for both groups of principals. Two

other generic conditions identified in the literature review included relevance and timeliness.

What did not appear in the literature was the importance of key values acting as catalytic conditions during Cyclical Reviews. The values related to human factors included trust, openness, credibility, competence, knowledge, commitment, skills, and ownership. The values related to evaluation factors were found to include openness, clarity, consistency, standards, consensus, engagement, transparency, and functionality. Finally, the context values were found to include openness, readiness, clarity, and transparency. While many of these values were recognised as characteristics or factors in the prior research, the analysis of data from this study led to the implication for leadership theory about Cyclical Reviews that these values act as threshold conditions for change.

This in turn leads to the first recommendation for theory that:

1. The sets of values in Table 8.1 be considered catalytic conditions for factors triggering influence during Cyclical Reviews.

Table 8.1 Catalytic Values Related to Factors Triggering Influence

Human factors	**Context factors**	**Evaluation factors**
openness	openness	openness
trust	readiness	clarity
credibility	clarity	consistency
competence	transparency	standards
knowledge		consensus
commitment		engagement
ownership		transparency
		functionality

While a full treatment of the role of catalytic values is beyond the scope of this study, as mentioned above, these findings suggest that the nature, distribution, and use of power in the context may mediate influence. A recent review of political ideologies in educational leadership (Macpherson, 2014) showed that a tradition of pragmatism has been gradually challenged by communitarianism, communicative rationalism, and egalitarian liberalism. The catalytic values in Table 8.1 are strongly indicative of influential leadership, with some indications of pragmatism, in a 'school as a community', given the valuing of competence, standards, and functionality. This implies that further research could use political philosophy to identify and evaluate how appropriately power is used to gain influence during school-based reviews that are cyclical.

The next issue regarding theory concerns the interplay between the three categories, as summarised in Chapter 6, Figure 6.4. The second recommendation for theory is therefore that:

2. A theory of school-based cyclical review acknowledge that the evaluation methodology used in Cyclical Reviews is primarily independent as a process but is differentially affected by context and human factors in each school setting.

Further to the above, a third recommendation for theory is that:

3. A theory of school-based cyclical review acknowledge that major source of variance to standardised review processes is likely to be the leadership services provided to the review team, the leadership by the host principal, and by the partnership between the review leader and the coach.

From the review of the literature, Cyclical Reviews were found to be a form of participatory evaluation with all of the features of PAR. ECB was found to be an organisational learning process intended to make routine quality evaluation concerned with continuous improvement. As a consequence of uncovering the very different experiences of Group 1 and Group 2 principals and the extent to which these

experiences impacted on evaluation practices and capacity building, the fourth recommendation is that:

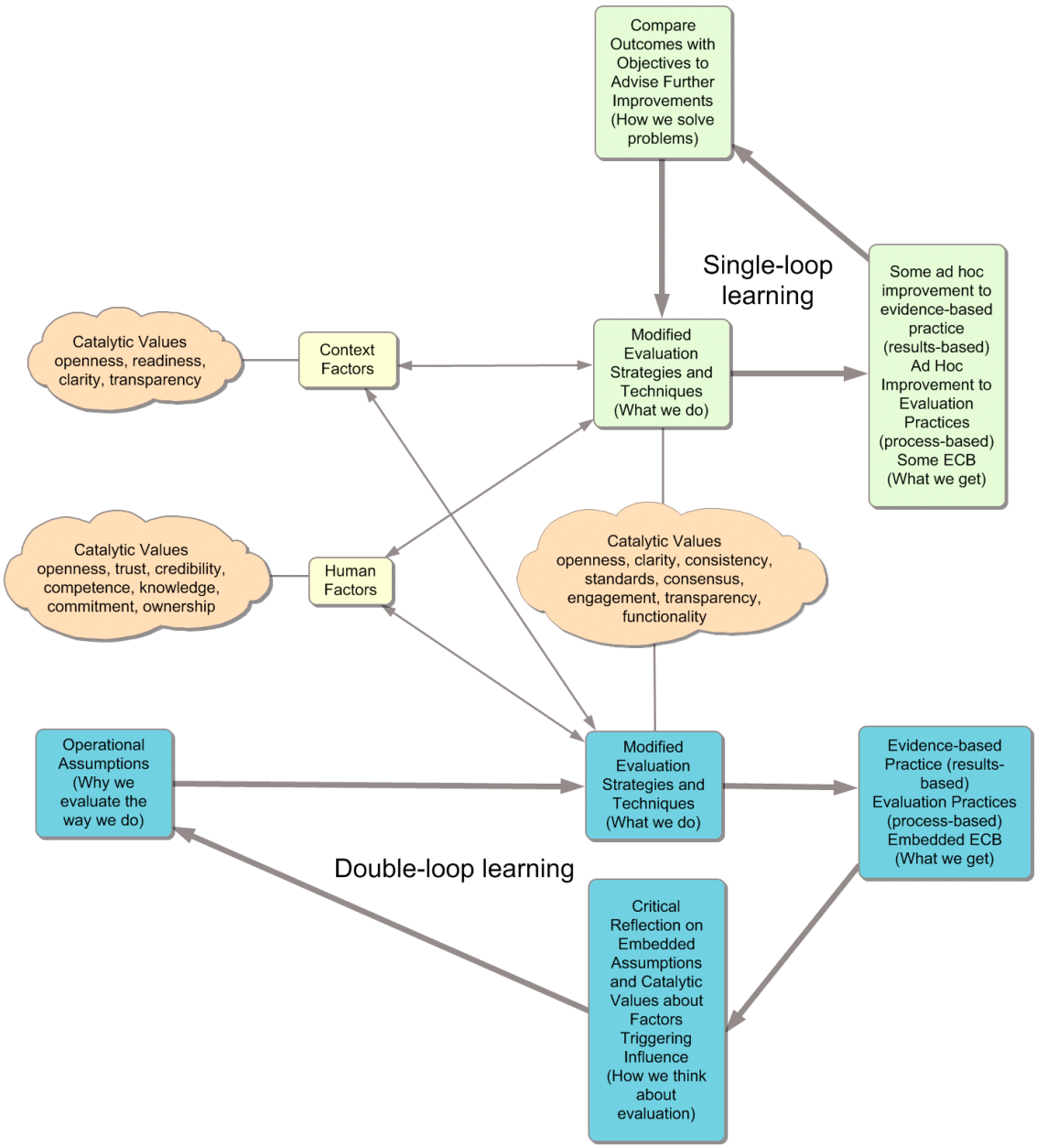
4. School-based reviews that are cyclical be theorised as PAR and organisational double-loop learning.

The recommendations have considerable strategic implications for school leaders in both their professional development and their use, development, and implementation of evaluation policy. The whole point of getting better value from the evaluations is to benefit key decision-making bodies, such as the Department, while at the same time to enhance the leadership capability and performance of the participating principals, whatever their role. Therefore the fifth recommendation for theory is that:

5. It be assumed that leaders of school-based reviews which are cyclical will need supervised professional development and leadership service to learn the theoretical and practical implications of double-loop learning.

The final elaborations on the basic model are presented in Figure 8.3 overleaf. Again, the diagram is shaded blue for double-loop learning, green for single-loop learning yellow for human and context factors (evaluation factors being subsumed in the double-loop and single-loop learning). In addition, the catalytic values are shown in an orange cloud.

Figure 8.3 Proposed Model of Evaluation Influence



Recommendations for Policy

Improvements to policy regarding Cyclical Reviews in New South Wales government schools were found in Chapters 6 and 7 to be warranted. The program reported here was confined to schools in the Western Sydney Region and this section make generalisations, cautiously, to state policy level.

The most important policy issue to emerge from this study was the need for Cyclical Reviews to retain all the features of PAR, which implied the need to organise reviews around cohorts of principals and for each future cohort to engage in experiential learning processes through designing, participating, and then reflecting critically on the quality of evaluation and the degree of capacity building achieved. This need also means that offering principals external validation services, even using other principals, does not necessarily provide them with the double-loop learning opportunities required to guarantee evidence-based evaluation practice and ECB. The first policy recommendation is that:

1. The Department develop a Cyclical Review policy with guidelines that cohere with PAR and double-loop learning principles and aim at sustaining ECB.

The second related policy recommendation is that:

2. The Department permit *Exemplary Practice Statements* and any national replacement of them to be considered as a basis for articulating customised evaluation criteria by each cohort at the commencement of its review process.

This approach would obviate the need for separate school accountability mechanisms in most cases, although 'reviews by exception' should remain available as a fail-safe device.

The third policy recommendation is that:

3. The Department formally implement a Cyclical Review Program, and that in consultation with principals the Department determine the purposes and infrastructure of that program in order to train and accredit team members, host principals, team leaders, and coaches.

Recommendations for Practice

From the analysis and synthesis of findings presented in Chapters 6 and 7 and in line with the research literature, the quality and relevance of evaluation practice, in all of its forms, was found to exert influence in three broad ways. First was the extent to which Cyclical Reviews were efficiently and effectively conducted. Second was the extent to which skills, understandings, and evaluation capacity of the principals participating in the reviews were developed. Third was the extent to which review findings and processes in the schools of the participating principals were implemented.

Practice in relation to planning processes was also found to be effective and efficient in that reviews were conducted on time, were completed following the planned procedures, and used the information-technology resources secured for the purpose. The review teams applied themselves with energy and conviction, and the schools under review appeared to have appreciated the process and the resulting recommendations for school and staff development. As also described, refinements to a number of the processes were made by the principals as they developed evaluation skills, knowledge, and understandings. Further, by time that the study was concluding advances in technologies had led the researcher and a number of participants to consider adopting newer, more efficient, and more effective data-gathering and sorting methods to either refine or replace PPODS. This is, again, consistent with double-loop learning.

A further area of practice was that of professional learning and accreditation. It was found that the principals not only acquired new evaluation skills and knowledge but recognised and valued these qualities in others and believed that the reviews were more effective when all involved—coaches, host principals, team leaders, and team members—were trained specifically for their roles.

A final area of practice concerned the composition and selection of review teams. It was found that the selection of staff or parents from the same school learning community could lead to cross-fertilisation of ideas and more cooperative planning between schools.

In light of these findings, four recommendations for practice that can now be made are that:

1. The current array of evaluation strategies and tools be retained and further improved by each cohort.
2. The current methods for data collection and PPODS be updated or replaced to exploit contemporary technologies.
3. Key competencies for each role—coach, team leader, host principal, and team member—be identified so that training lead to differentiated accreditation.
4. Criteria and processes for selecting participation by principals, school staff, and parent and community members in Cyclical Review teams be determined.

Recommendations for Further Research

This study indicated a number of potential areas for research that can be expressed as research questions:

1. To what extent does single-loop and double-loop learning serve as an explanatory heuristic from a systematic review of school review processes internationally?
2. What further links can be found between ECB, self-evaluation, OD, and organisational learning?
3. Using political philosophy, how appropriately is power used to gain influence during school-based reviews which are cyclical?
4. What is the relative impact of Cyclical Reviews on student learning compared with other forms of school evaluation and accountability?
5. Are there further values that serve as catalytic conditions for factors triggering influence during Cyclical Reviews?

6. To what extent does the proposed model of evaluation influence explain the outcomes of future Cyclical Reviews and how can they be further improved?
7. How do Cyclical Reviews influence participants other than principals?
8. What is the relative extent of ECB on principals in their respective roles of team leaders, host principals, or both?

CONCLUSION

The influence that the Cyclical Review process had on participating school principals varied with involvement. Those principals who were fully involved in the design and application of, and critical reflection on, the evaluation processes engaged in deep learning, consistent with double-loop learning. Catalytic values mediated the impact of context, human, and evaluation factors. Theory, policy, practice, and further research about school-based cyclical reviews should all be refined to reflect the principles of double-loop learning.

REFERENCES

Adler, P. A., & Adler, P. (1994). Observational techniques. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 377–392). Thousand Oaks, CA: Sage.

Ainscow, M. (2005). *The next big challenge: inclusive school improvement*. Paper presented at the International Congress of School Effectiveness and Improvement, Barcelona, January, 2005, 10.

Alkin, M. C. (1985). *A Guide for Evaluation Decision Makers*. Beverly Hills, CA: Sage.

Alkin, M. C. (2003). Introduction (to section on Evaluation Use). In T. Kelligan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (pp. 189–196). Dordrecht: Kluwer Academic Press.

Alkin, M. C. (2004a). Comparing evaluation points of view. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: Sage.

Alkin, M. C. (Ed.). (2004b). *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks: Sage.

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 12–65). Thousand Oaks, CA: Sage.

Alkin, M. C., Daillak, R., & White, P. (1979). *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills: Sage.

Alkin, M. C., Koescoff, J., Fitz-Gibbon, C., & Seligman, R. (1974). *Evaluation and Decision Making: The Title VII Experience*. Los Angeles: Center for the Study of Evaluation.

Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation, 29*, 1–12.

Angus, L. B. (1987). A critical ethnography of continuity and change in a catholic school. In R. J. S. Macpherson (Ed.), *Ways and Meaning of Research in Educational Administration*. Armidale: University of New England.

Argyris, C. (1996). *On Organizational Learning*. Cambridge, MA: Blackwell.

Argyris, C., & Schön, D. A. (1978). *Organizational Learning: A theory of action perspective*. Reading MA: Addison-Wesley.

Argyris, C., & Schön, D. A. (1996). *Organizational Learning II: Theory, Method and Practice*. Reading, MA: Addison-Wesley.

Australian Capital Territory Department of Education and Training. (2009). School Improvement Framework. Better schools... better futures. Raising Quality and Achieving Excellence in ACT Public Schools Retrieved 2 November 2009, from http://www.det.act.gov.au/publications_and_policies/publications_a-z/school_improvement

Australian Concise Oxford Dictionary. (1997) (3rd ed.). Melbourne: Oxford University Press.

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002a). Editors' Note. *New Directions for Evaluation: The Art, Craft, and Science of Evaluation Capacity Building, 93*, 1–6.

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002b). New Directions for ECB. *New Directions for Evaluation: The Art, Craft, and Science of Evaluation Capacity Building, 93*, 109–119.

Bannister, D., & Fransella, F. (1971). *Inquiring Man: The Theory of Personal Constructs*. Harmondsworth: Penguin.

Bannister, D., & Fransella, F. (1986). *Inquiring Man: The Psychology of Personal Constructs* (3rd ed.). London: Taylor and Francis Group.

Bannister, D., & Mair, J. M. M. (1968). *The Evaluation of Self Constructs*. London: Academic Press.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, NJ: Prentice-Hall.

Boston, K. G. B. (1999). *Adult education and lifelong learning.* Paper presented at the Adult Education Conference, Newcastle.

Bourgeios, I., & Cousins, J. B. (2013). Understanding Dimensions of Organizational Evaluation Capacity. *American Journal of Evaluation, 34*(3), 299–319.

Boyle, R., Lemaire, D., & Rist, R. C. (1999). Introduction: Building evaluation capacity. In R. Boyle & D. Lemaire (Eds.), *Building Effective Evaluation Capacity: Lessons from Practice* (pp. 1–22). New Brunswick, NJ: Transaction Pubishers.

Brandt, R. M. (1982). The new catechism for effective schools. *Educational Leadership, December*, 23.

Brydon-Miller, M., Kral, M., Maguire, P., Noffke, S., & Sabhlok, A. (2011). Jazz and the Banyan Tree: Roots and Riffs on Participatory Action Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research* (pp. 387–400). Los Angeles, CA: Sage.

Caracelli, V. J., & Preskill, H. (2000). *New Directions for Evaluation* (Vol. 88). San Francisco: Jossey-Bass.

Carnoy, M., Elmore, R., & Sisken, L. S. (Eds.). (2003). *The New Accountability*. New York: RoutledgeFalmer.

Charmanz, K. (2006). *Constructing Grounded Theory*. Thousand Oaks, CA: Sage.

Christie, C. A. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation, 28*, 8–25.

Ciarlo, J. A. (1981). Editor's Introduction. In J. A. Ciarlo (Ed.), *Utilizing Evaluation: Concepts and Measurement Techniques*. Thousand Oaks, CA: Sage.

Coffey, A., & Atkinson, P. (1996). *Making Sense of Qualitative Data: Complementary Research Strategies*. Thousand Oaks, CA: Sage.

Cohen, L., Manion, L., & Morrison, K. R. B. (2007). *Research Methods in Education* (6th ed.). London: Routledge.

Compton, D. W., & Baizerman, M. (2007). Defining evaluation capacity building. *American Journal of Evaluation*, *28*(1), 118–119.

Compton, D. W., Glover-Kudon, R., Smith, I. E., & Avert, M. E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995–2001. *New Directions for Evaluation, 93*, 47–61.

Connell, R., & Campbell, C. (2007). Researching education. In R. Connell, Campbell, M. Vickers, A. Foley, D. and Bagnall, N. (Ed.), *Education, Change and Society* (pp. 301–325).

Cooksey, R., & McDonald, G. (2011). *Surviving and Thriving in Postgraduate Research*. Prahan: Tilde University Press.

Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, D. L. Stufflebeam & L. A. Wingate (Eds.), *International Handbook of Educational Evaluation* (pp. 245–265). Boston: Kluwer.

Cousins, J. B., Goh, S. C., Clark, S., & Lee, L. E. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *The Canadian Journal of Program Evaluation, 19*(2), 99–141.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56*(3), 331–364.

Cousins, J. B., & Leithwood, K. A. (1993). Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: Creation, Utilization, Diffusion, 14*(3), 305–333.

Cousins, J. B., & Shulha, L. M. (2006). A comparative analysis of evaluation utilization and its cognate fields of inquiry: Current issues and trends. In I. F. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of Evaluation: Policies, Programs and Practices*. London, Thousand Oaks, CA: Sage.

Cousins, J. B., & Whitmore, E. (2007). Framing participatory evaluation. *New Directions for Evaluation, 80, 1998, reprinted in Enduring Issues in Evaluation: The 20th Anniversary of the Collaboration between NDE and AEA, New Directions for Evaluation, 114*, 87–105.

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory in Practice, 39*(3), 124–130.

Cronbach, L. J., & Associates. (1980). *Toward Reform of Program Evaluation*. San Francisco: Jossey-Boss.

Cummings, R. (1999). *Stakeholder involvement: a foundation for evaluation use*. Paper presented at the American Evaluation Society Conference, Orlando, Florida, Orlando Florida.

Cummings, R. (2002). *Rethinking Evaluation Use*. Paper presented at the Australasian Evaluation Society International Conference, Wollongong, Wollongong.

Cummings, T. (2008). *Organization Development and Change* (9th ed.). New York, New York: Thomson Wardsworth.

Dabelstein, N. (2003). Evaluation capacity development: Lessons learned. *Evaluation, 9*(3), 365.

Dalton, P., & Dunnett, G. (1992). *A Psychology for Living: Personal Construct Theory for Professionals and Clients*. Chichester, England; New York: J. Whiley and Sons.

De Grauwe, A., & Naidoo, J. (2004). *School evaluation for quality improvement*. Paper presented at the Meeting of the Asian Network of Training and Research Institutions in Educational Planning, 2–4 July 2002, Kuala Lumpur, Malaysia, published in UNESCO: International Institute for Educational Planning, Paris, 2004 Kuala Lumpur, Malaysia.

De Walt, K. M., & DeWalt, B. R. (2002). *Participant Observation: A Guide for Fieldworkers*: Rowman Altamira.

Degenhardt, L., & Duignan, P. (2010). *Dancing on a Shifting Carpet: Reinventing Traditional Schooling for the 21st Century*. Camberwell, Vic.: ACER Press.

Degenhardt, L., & McCulla, N. (2013). *Who'd be a School Principal Anyway?* Paper presented at the School Leadership Symposium, Zug, Switzerland, 26–28 September 2013. http://www.edulead.net/2013/speakers.php

Delaney, T. A. (2006). *Evaluation influence and the most significant change (MSC) technique*. Paper presented at the 2006 Australasian Evaluation Society International Conference, Darwin, Australia, 5-7 September 2006.

Denzin, N. K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods* (2nd ed.). Chicago: Aldine.

Denzin, N. K., & Lincoln, Y. S. (2008a). Introduction: The Discipline and Practice of Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Inquiry* (3rd ed.). Thousand Oaks: Sage.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2008b). *Strategies of Qualitative Inquiry* (3rd ed.). Thousand Oaks: Sage.

Department of Education and Children's Services. (2007). DECS Improvement and Accountability Framework Retrieved 3 August 2009, from http://www.decs.sa.gov.au/

Diesing, P. (1972). *Patterns of Discovery in Social Sciences*. London: Routledge and Keegan Paul.

Dolan, S. L., Garcia, S., & Richley, B. (2006). *Managing by Values: A Corporate Guide to Living, Being Alive, and Making a Living in the 21st Century*. London: Palgrave Macmillan.

Dugdale, J. H. (1990). *Inspectors of schools in the state education system of New South Wales, 1848–1971.* (PhD), University of Sydney.

Dugdale, J. H. (2005). A brief history of the public school inspection system in New South Wales. In L. Craddock (Ed.), *I Remember When. . .* (pp. 4-9): Association of Retired Inspectors of Schools and Senior Educational Administrators.

Education Queensland. (2009). *Queensland School Improvement and Accountability Framework*. Brisbane: Education Queensland.

Feagin, J. R., Orum, A. M., & Sjoberg, G. (Eds.). (1991). *A case for the case study*. Chapel Hill, NC: University of North Carolina Press.

Fielding, N. (1996a). Ethnography. In N. Gilbert (Ed.), *Researching Social Life* (pp. 154–171). London: Sage.

Fielding, N. (1996b). Qualitative interviewing. In N. Gilbert (Ed.), *Researching Social Life* (pp. 154–171). London: Sage.

Finnegan, R. (1996). Using documents. In R. Sapsford & V. Jupp (Eds.), *Data Collection and Analysis* (pp. 138–151). London: Sage.

Forss, K., Rebien, C. C., & Carlsson, J. (2002). Process use of evaluation: Types of use that precede lessons learned and feedback. *Evaluation, 8*(1), 29–45.

Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison*. Harmondsworth: Penguin.

Fowler, H. W. (1965). *A Dictionary of Modern English Usage* (2nd ed.). Oxford: Oxford University Press.

Fransella, F. (1975). *Need to Change?* London: Methuen.

Fullan, M. (2001). *The New Meaning of Educational Change* (3rd ed.). New York: Teachers College.

Giddens, A. (1984). *The Constitution of Society: Outline of a Theory of Structuration*. Berkeley and Los Angeles: University of California Press.

Gilliam, A., Barrington, T., Davis, D., Lacson, R., Uhl, G., & Pheonix, U. (2003). Building evaluation capacity for HIV prevention programs. *Evaluation and Program Planning, 26*, 133–142.

Ginsberg, A., & Rhett, N. (2003). Building a Better Body of Evidence: New Opportunities to Strengthen Evaluation Utilization *American Journal of Evaluation, 24*(4), 489–498.

Glaser, B., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.

Gold, R. L. (1958). Roles in Sociological Field Observations *Social Forces, 36*(3), 217–223.

Greene, J. C. (1980). Communication of results and utilization in participatory program evaluation. *Evaluation and Program Planning, 11*, 331–351.

Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12*(2), 341–351.

Greene, J. C. (2004). The educative evaluator: An interpretation of Lee. J. Cronbach's vision of evaluation. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 169–180). Thousand Oaks: Sage.

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

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105–117). Thousand Oaks, CA: Sage.

Hammersley, M., & Atkinson, P. (1995). *Ethnography: Principles in Practice* (2nd ed.). London: Routledge.

Hargreaves, A. (2013). *5 I's of Educational Change*. Paper presented at the School Leadership Symposium 2013, Zug, Switzerland, 26–28 September 2013, Zug, Switzerland. http://www.edulead.net/2013/speakers.php

Hattie, J. (2005). *What is the nature of evidence that makes a difference to learning?* Paper presented at the Using Data to Support Learning ACER Research Conference, 7–9 August, 2005, Melbourne, Victoria, published in Conference Proceedings, Australian Council for Educational Research,12–13, Melbourne, Victoria.

Helmstadler, G. C. (1970). *Research Concepts in Human Behavior*. New York: Appleton-Century-Crofts.

Henry, G. T. (2000). Why not use? In V. J. Caracelli & H. Preskill (Eds.), *The Expanding Scope of Evaluation Use. New Directions for Evaluation* (Vol. 88, pp. 85–98). San Francisco: Jossey-Bass.

Henry, G. T. (2003). Influential evaluations. *American Journal of Evaluation, 24*(4), 515–524.

Henry, G. T., & Mark, M. M. (2003). Beyond use: understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation, 24*(3), 293–314.

Hofstetter, C. H., & Alkin, M. C. (2003). Evaluation use revisited. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Press.

Huberman, A. M. (1987). Steps towards an integrated model of research utilization. *Knowledge: Creation, Utilization, Diffusion, 8*(4), 586–611.

Ikin, K. B. (2006). *Performance Agreement*. Western Sydney Region, Department of Education and Training. Kingswood.

Ikin, K. B. (2007). *Plan to support the implementation of Cyclical Reviews in Western Sydney Region*. submission to the Regional Director. Western Sydney Region New South Wales Department of Education and Training.

Imants, J. (2003). Two basic mechanisms for organizational learning in schools. *European Journal of Teacher Education, 26*(3), 293-311.

Ingram, G. K., Fostvedt, N., Feinstein, O., & Mackay, K. (2002). 2002 Annual Report on Evaluation Capacity Development Retrieved from www.worldbank.org/oed

Jeris, L. (2003). Double-loop Learning in Teams. *Lifelong Learning in Europe, 8*, 8–17.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377–410.

Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning, 21*(1), 93–110.

Joint Committee on Standards for Educational Evaluation. (1994). *The Program Evaluation Standards* (2nd ed.). Thousand Oaks, CA: Sage.

Jupp, V. (1996). Documents and critical research. In R. Sapsford & V. Jupp (Eds.), *Data Collection and Analysis* (pp. 283–316). London: Sage.

Kelly, G. A. (1955). *The Psychology of Personal Constructs* (Vol. 1 and 2). New York: Norton.

Kelly, G. A. (1970). A brief introduction to personal construct theory. In D. Bannister (Ed.), *Perspectives of Personal Construct Theory*. New York: Academic Press.

Kemmis, S., & McTaggart, R. (2008). Participatory Action Research: Communicative Action and the Public Sphere. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Inquiry* (pp. 271–330). Thousand Oaks, CA: Sage.

King, J. A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation, 93*, 63–80.

King, J. A. (2007). Making sense of participatory evaluation. *Enduring Issues in Evaluation: The 20th Anniversary of the Collaboration between NDE and AEA, New Directions for Evaluation, 114*, 83–86.

King, J. A., & Pechman, E. M. (1984). Pinning a wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis, 6*, 241–451.

King, J. A., & Thompson, B. (1983). Research on school use of program evaluation: A literature review and research agenda. *Studies in Educational Evaluation, 9*, 5–21.

King, J. A., & Volkov, B. (2005). A Framework for Building Evaluation Capacity Based on the Experiences of Three Organizations. *CURA Reporter, 35*(3), 10–16.

Kirkhart, K. (2000). Reconceptualizing evaluation use: an integrated theory of influence. In V. J. Caracelli & H. Preskill (Eds.), *The Expanding Scope of Evaluation Use. New Directions for Evaluation* (Vol. 88, pp. 5–24). San Francisco: Jossey-Bass.

Kraemer, H. M. J. (2011). *From Values to Action: The Four Principles of Values-Based Leadership*. San Francisco: Jossey-Bass.

Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. Newbury Park, CA: Sage.

Larsen, J. K., & Werner, P. D. (1981). Measuring utilization of mental health program consultation. In J. A. Ciarlo (Ed.), *Utilizing Evaluation: Concepts and Measurement Techniques* (pp. 77–96). Beverly Hills, CA: Sage.

Lawrenz, F., Gullickson, A., & Toal, S. A. (2007). Dissemination: Handmaiden to evaluation use. *American Journal of Evaluation, 28*(3), 275–289.

Lawrenz, F., King, J. A., & Greenseid, L. O. (2005). *The relationship between involvement and use/influence in large multi-site evaluations*. Paper presented at the joint meeting of the Canadian Evaluation Society and the American Evaluation Association, Toronto, Canada.

LeCompte, M. D., Preissle, J., & Tesch, R. (1993). *Ethnography and Qualitative Design in Educational Research*. San Diego, CA: Academic Press.

Leviton, L. C. (2003). Evaluation use: advances, challenges and applications. *American Journal of Evaluation, 24*(4), 525–535.

Leviton, L. C., & Hughes, E. F. X. (1981). Research on the utilization of evaluations: A review and synthesis. *5, 4*(525–548).

Linn, R. (2000). Assessment and accountability. *Educational Researcher, 25*, 4–16.

MacBeath, J. (1999). *Schools Must Speak for Themselves*. London: RoutledgeFalmer.

MacBeath, J. E. C., & McGlynn, A. (2002). *Self-evaluation: What's in it for Schools?* London: Routledge.

MacDonald, K., & Tipton, C. (1996). Using documents. In N. Gilbert (Ed.), *Researching Social Life* (pp. 187–200). London: Sage.

MacLellan-Wright, M. F., Patten, S., dela Cruz, A. M., & Flaherty, A. (2007). A participatory approach to the development of an evaluation framework: Process, pitfalls, and payoffs. *Canadian Journal of Program Evaluation, 22*(1), 99–124.

Macpherson, R. J. S. (1982). The 'I and We' of Accountability: An Example of Meta-Evaluation in Education Retrieved 12 January 2007, from Eric document retrieval services ED 229 837

Macpherson, R. J. S. (1983). The Western Australian peer process consultancy project: Action research as INSET for principals. *British Journal of Inservice Education, 9*(3), 141–149.

Macpherson, R. J. S. (2014). *Political Philosophy, Educational Administration and Educative Leadership*. London: Routledge.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation, 10*(1), 35–57. Retrieved from http://evi.sagepub.com/cgi/content/abstract/10/1/35 doi:10.1177/1356389004042326

McCall, G. J., & Simmons, J. L. (Eds.). (1969). *Issues in Participant Observation: A Text and Reader*. Massachusetts: Addison-Wesley.

McClenaghan, P. B. (2006). *How Australian Vice-Chancellors Do Their Jobs—A Role-Theory Analysis*. Doctoral Thesis. UNE.

McCoy, A., Rose, D., & Connolly, M. (2013). Developing Evaluation Cultures in Human Service Organisations. *Evaluation Journal of Australia, 13*(1), 15–20.

McDonald, B., Rogers, P., & Kefford, B. (2003). Teaching people to fish? Building the evaluation capability of public sector organizations. *Evaluation, 9*(1), 9–29.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Miller, W. L., & Crabtree, B. F. (2008). Clinical Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Inquiry* (pp. 331–376). Thousand Oaks, CA: Sage.

Milstein, B., Chapel, T. J., Wetterhall, S. F., & Cotton, D. A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation, 93*, 27–46.

Minichiello, V., Aroni, R., Timewell, E., & Alexander, L. (1990). *In-depth Interviewing: Researching People*. Melbourne: Longman Cheshire.

Morgan, C. (1980). Paradigms and metaphors and puzzle solving in organisational theory. *Administrative Science Quarterly, 25*(4), 605–622.

Mott, A. (2003). Hand in hand—evaluation and organizational development. *The Evaluation Exchange, 9*(3), 8.

Muscatello, D. B. (1988). Developing an agenda that works: The right choice at the right time. *New Directions for Program Evaluation, 39*, 21–31.

Nevo, D. (1995). *School-based Evaluation: A Dialogue for School Improvement*. Oxford: Pergamon.

Nevo, D. (2006). Evaluation in education. In I. F. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of Evaluation*. London: Sage.

New South Wales Department of Education and Training. (1999). *School Development Policy*. Sydney: New South Wales Department of Education and Training.

New South Wales Department of Education and Training. (2003). *Lifelong Learning: The Future of Public Education in New South Wales*. Sydney: New South Wales Department of Education and Training.

New South Wales Department of Education and Training. (2004). *School Development Policy*: New South Wales Department of Education and Training.

New South Wales Department of Education and Training. (2005). Framework for School Development & Accountability Retrieved 20 February 2006, from https://detwww.det.nsw.edu.au/directorates/schoimpro/sda//sda.htm

New South Wales Department of Education and Training. (2006). *School Development Policy: School Review Resources*: New South Wales Department of Education and Training.

New South Wales Department of Education and Training Western Sydney Region. (2008). *Submission and related papers for the establishment of the Cyclical Review Steering Committee*.

New South Wales Department of Education and Training Western Sydney Region. (February 2008). *Cyclical Reviews Steering Committee Meeting Minutes*.

New South Wales Department of Education and Training Western Sydney Region. (June 2006). *Cyclical Review Working Party Meeting Minutes*. Kingswood.

New South Wales Department of Education and Training Western Sydney Region. (September 2008). *Cyclical Review Steering Committee Meeting Minutes*.

New South Wales Department of School Education. (1992). *Strategic Plan 1993–1997*. Market Street, Sydney: New South Wales Department of School Education.

New South Wales Department of School Education. (1997). *Annual Report 1997*. Sydney: New South Wales Department of School Education.

New South Wales Department of School Education. (n.d.-a). *School Accountability and Improvement in NSW Public Schools*. Sydney: New South Wales Department of School Education.

New South Wales Department of School Education. (n.d.-b). *School Accountability and Improvement in NSW Public Schools: Rationale*. Sydney: New South Wales Department of School Education.

Northern Territory Department of Employment Education and Training. (2006). DEET School Planning and Accountability Framework Retrieved 2 May 2006, from http://www.deet.nt.gov.au/education

Northern Territory Department of Employment Education and Training. (2009). DEET Accountability and Performance Improvement Framework Guide for Schools Retrieved 2 November 2009, from https://www.det.nt.gov.au/about-us/APIF

Oppenheim, A. N. (2001). *Questionnaire design, interviewing and attitude measurement* (2nd ed.). London: Continuum International Publishing Group.

Owen, J. M. (2006). *Program Evaluation: Forms and Approaches* (3rd ed.). Crows Nest, Australia: Allen and Unwin.

Patton, M. Q. (1978). *Utilization-focused Evaluation*. Beverly Hills, CA: Sage.

Patton, M. Q. (1980). *Qualitative Evaluation Methods*. Newbury Park, CA: Sage.

Patton, M. Q. (1997). *Utilization-focused Evaluation: The New Century Text* (3rd ed.). Thousand Oaks, CA: Sage.

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

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making* (pp. 141–163). Lexington MA: Lexington Books.

Pelto, P. J., & Pelto, G. H. (1976). *Anthropological Research: The Structure of Inquiry* (2nd ed.). Cambridge: Cambridge University Press.

Perryman, J. (2005). School leadership and management after special measures: discipline without the gaze? *School Leadership and Management, 25*(3), 281–297.

Pope, A. M. (1980). The teacher as learner—some factors in the learning process. *Britain Journal of In-Service Education, 7*(1), 70–76.

Preskill, H. (1991). The cultural lens: Bringing utilization into focus. *New Directions for Program Evaluation, 49*, 5–15.

Preskill, H., & Caracelli, V. J. (1997). Current and developing conception of use: evaluation using TIG survey results. *Evaluation Practice, 18*(3), 209–226.

Preskill, H., & Torres, R. T. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: Sage.

Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation, 24*(4), 423–442.

Punch, K. F. (2009). *Introduction to Research Methods in Education*. London: Sage.

Radii. (2005). Planning and Accountability for School Improvement: National and International Literature Review. Retrieved from

Rich, R. F. (1977). Use of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C. H. Weiss (Ed.), *Using Social Science Research in Public Policy Making* (pp. 199–211). Lexington, MA: Lexington-Heath.

Rist, R. C. (1977). On the relations among educational research paradigms: disdain to detente. *Anthropology and Education Quarterly, 13*(2), 42–49.

Rogers, C. R. (1964). Towards a science of personality. In T. W. Mann (Ed.), *Behaviourism and Phenomenology*. Chicago: Chicago University Press.

Rossman, G., & Rallis, S. (2000). Critical inquiry and use in action. In V. J. Caracelli & H. Preskill (Eds.), *The Expanding Scope of Evaluation Use. New Directions in Evaluation* (Vol. 88, pp. 55–69). San Fransisco, CA: Jossey-Bass.

Russ-Eft, D., & Preskill, H. (2001). *Evaluation in Organization: A Systematic Approach to Enhance Learning, Performance, and Change*. Cambridge, MA: Perseus Publishing.

Ryan, G. W., & Bernard, H. R. (2003a). Data Management and Analysis Methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (2nd ed.). Thousand Oaks: Sage.

Ryan, G. W., & Bernard, H. R. (2003b). Techniques to Identify Themes. *Field Methods, 15*(1), 85–109.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key Characteristics of Effective Schools; A Review of School Effectiveness Research*. London: Institute of Education and Ofsted.

Scharmer, C. O. (2009). *Theory U: Leading from the Future as it Emerges*. San Francisco: Berrett-Koehler Publishers.

Scharmer, C. O., & Kaufer, K. (2013). *Leading from the Emerging Future: From Ego-System to Eco-System Economics*. San Francisco: Berrett-Koehler Publishers.

Schatzman, L., & Strauss, A. L. (1973). *Field Research: Strategies for a Natural Sociology*. New Jersey: Prentice Hall.

Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). Thousand Oaks, CA: Sage.

Senge, P. M. (2010). The Fifth Discipline: The Art & Practice of the Learning Organization (revised ed.). London: Crown Publishing Group.

Senge, P. M., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that Learn*. New York: Currency/Doubleday.

Senge, P. M., Scharmer, C. O., Jaworski, J., & Flowers, B. S. (2008). *Presence: Human Purpose and the Field of the Future*. London: Crown Business.

Senge, P. M., Smith, B., Kruschwitz, N., Laur, J., & Schley, S. (2010). *The Necessary Revolution: How Individuals And Organizations Are Working Together to Create a Sustainable World*. New York: Broadway Books.

Sergiovanni, T. J. (1977). *Handbook for Effective Department Leadership*. Boston: Allyn and Bacon.

Shadish, W., Cook, T., & Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *American Journal of Evaluation, 18*(3), 195–208.

Silverman, D. (1993). *Interpreting Qualitative Data: Methods for Analyzing Talk, Text, and Interaction*. London: Sage.

Smith, L. M. (1982). Ethnography. In H. E. Mitzel, J. H. Best & W. Rabonowitz (Eds.), *The Encyclopedia of Educational Research* (5th ed., pp. 587–591). New York, New York: Free Press.

Spradley, J. P. (1980). *Participant Observation*. New York: Holt, Rheinhart and Winston.

Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 236–247). Thousand Oaks, CA: Sage.

Stake, R. E. (2006). *Multiple Case Study Analysis*. New York: Guilford Press.

Stern, P. C. (1979). *Evaluating Social Science Research*. Oxford, New York: OUP.

Stoll, L. (2013). *Creativity in Leadership: Issues and Possibilities*. Paper presented at the School Leadership Symposium 2013, Zug, Switzerland, 26–28 September 2013, Zug, Switzerland. http://www.edulead.net/2013/speakers.php

Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation, 89*, 7–98.

Stufflebeam, D. L. (2002). Institutionalizing evaluation checklist. Retrieved from http://www.wmich.edu/evalctr/checklists/institutionalizingeval.htm.

Stufflebeam, D. L. (2004). The 21st-century CIPP model: Origins, development, and use. In M. C. Alchin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 245–266). Thousand Oaks, CA: Sage.

Sudman, S. B., & Bradburn, N. M. (1982). *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

Tasmanian Department of Education and Training. (2009). Department of Education Retrieved 2 November 2009, from http://www.education.tas.gov.au/

Taut, S. (2007). Defining evaluation capacity building: Utility considerations. *American Journal of Evaluation, 28*(1), 120.

Tesch, R. (1990). *Qualitative Research: Analysis Types and Software Tools*. Basingstoke: Falmer.

Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation, 22*, 387–395.

Tyler, R. W. (1950). *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago.

Victorian Department of Education and Training. (2006). Accountability and Improvement Framework for Victorian Government Schools 2007 Retrieved 1 March 2007, from http://www.education.vic.gov.au/management/schoolimprovement/default.htm

Vilkinas, T., & Ladyshewsky, R. K. (2011). Leadership behaviour and effectiveness of academic program directors in Australian universities. *Educational Management Administration & Leadership, 40*(1), 109–126.

Wasson, L. (2006a, 1 March 2006). [Meeting with School Development Officers].

Wasson, L. (2006b, 15 and 22 November 2006). [Meetings with Principals' Groups].

Wasson, L. (2007, 16 January 2007). [Meeting with school development officers].

Wasson, L. (2008, 12 February). [Letter of invitation to Steering Committee Members].

Weiss, C. H. (1973). Where politics and evaluation research meet. *Evaluation, 1*(3), 37–45.

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review, 39*, 426–431.

Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Utilization, Diffusion, 1*(3), 381–404.

Weiss, C. H. (1981). Measuring the use of evaluation. In J. A. Ciarlo (Ed.), *Utilizing Evaluation: Concepts and Measurement Techniques* (pp. 17–33). Beverly Hills, CA: Sage.

Weiss, C. H. (1982a). Policy research in the context of diffuse decision making. *The Journal of Higher Education, 53* (6), 619–639.

Weiss, C. H. (1982b). Measuring the use of evaluation. *Evaluation Studies Review Annual*, 129–145.

Weiss, C. H. (1998a). *Evaluation Research: Methods for Studying Program and Policies* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

Weiss, C. H. (1998b). Improving the use of evaluations: Whose job is it anyway? *Advances in Educational Productivity, 7*, 263–276.

Weiss, C. H. (Ed.). (1977). *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternative route to policy influence: how evaluations affect D.A.R.E. *American Journal of Evaluation, 26*(1), 12–30.

Western Australia Department of Education and Training. (2006). School Accountability Framework Review: National and International Perspectives and Approaches: Research Papers – Executive Summary (December 2006) Retrieved 1 March 2007, from http://www.det.wa.edu.au/education/accountability

Western Australia Department of Education and Training. (2008). School Improvement and Accountability Framework Retrieved 3 June 2008, from http://www.det.wa.edu.au/education/accountability/

Western Sydney Region New South Wales Department of Education and Training. (2005). *Western Sydney Region Strategic Plan 2006–2008*: Western Sydney Region New South Wales Department of Education and Training.

Western Sydney Region New South Wales Department of Education and Training. (2006). *Exemplary Practice Statements*. working document. Western Sydney Region. Kingswood.

Whitehead, I. (2005). Doing the diagnostic. *Prime Focus: The Professional Journal for Australian Primary School Leaders, 40*(Autumn), 2–4.

Wiersma, W., & Jurs, S. G. (2005). *Research Methods in Education: An Introduction* (8th ed.). Boston: Pearson/ A and B.

Wolcott, H. E. (1988). Ethnographic research in education. In R. M. Jaeger (Ed.), *Complementary Methods for Research in Education* (pp. 187–210). Washington DC: American Educational Research Association.

Yin, R. K. (2003). *Case Study Research: Design and Methods* (Third ed. Vol. 5). Thousand Oaks: Sage.