



MINISTRY OF EDUCATION

Te Tāhuhu o te Mātauranga

New Zealand

CONTRACT RESEARCH

Strategic Research Initiative

Literature Review

The Effects of Curricula and Assessment on Pedagogical Approaches and on Educational Outcomes

Report to the Ministry of Education

M. Carr, C. McGee, A. Jones, E. McKinley, B. Bell, H. Barr, T. Simpson

RESEARCH DIVISION

Wāhanga Mahi Rangahau

Authors: Malcolm Carr, Clive McGee, Alister Jones, Elizabeth McKinley, Beverley Bell,
Hugh Barr and Tina Simpson

Prepared for the New Zealand Ministry of Education under contract between the Ministry
and the University of Waikato

Executive Summary

The literature search was largely confined to large empirical studies and meta-analyses carried out in the last 5 years. The process adopted used electronic and physical searches and consultation with national and international experts. Around 200 references are referred to in the body of the review. Approximately 400 other references were checked for relevance. To assist the analysis a framework was developed which showed the interactions between curriculum, assessment and pedagogy, and the intervening role of the teacher and learner between this interaction and learning outcomes.

Key findings:

- The review identified the complex and interactive nature of the teaching and learning process, and an attempt was made to describe effective learning.
- Little research evidence was found for the effectiveness of different curriculum structures, despite widespread reforms. Some evidence for the value of clear curriculum structures was found. Aspects of assessment processes conflict with curriculum intentions.
- Strong evidence of the value of formative assessment to enhance learning was identified. A tension was found between formative and summative assessment processes which strongly influences learning achievement.
- Teaching is a complex activity which is strongly affected by external decisions about curriculum and assessment. The teacher plays a crucial role in the implementation of curriculum and assessment, and in student achievement. Research has identified some teaching effectiveness factors, e.g teacher expectations of their students, teacher knowledge of their subject, and other learning approaches which contribute to the feedback associated with formative assessment strategies.
- The report identifies significant gaps in the research on curriculum, assessment and pedagogy. Despite many innovations in these areas there is a dearth of research evidence about what happens in the classroom to the individual learner. This evidence is crucial to identifying innovations to enhance learning. Chapter 8 provides detail.

Table of Contents

Executive Summary.....	4
Table of Contents.....	5
Chapter 1: Overview.....	7
Chapter 2: The review process and outcomes.....	9
Introduction.....	9
2.1 The review requirements.....	9
2.2 The process of the review.....	10
2.3 Issues related to the complexity of the review process.....	13
Chapter 3: Towards a framework for analysis.....	17
Introduction.....	17
3.1 Effective learning.....	17
3.2 The intervening role of the teacher.....	20
3.3 The role of the student.....	21
3.4 Influences from outside the school.....	22
3.5 A framework for the review analysis.....	22
Chapter 4: Curriculum.....	25
Introduction.....	25
4.1 Commentary on curriculum.....	25
4.2 Research on Curriculum.....	26
4.2.1 Should curriculum be mandated or open?.....	26
4.2.2 What should be in the curriculum?.....	27
4.2.2.1 Subject curriculum (including the provision of differing curricula for different groups in the school system),.....	27
4.2.2.2 The relative importance of subjects.....	34
4.2.2.3 New subjects in the curriculum.....	35
4.2.2.4 The status of different features of a curriculum.....	35
4.2.3 The relationship between curriculum and assessment.....	35
4.3 The effects of Curriculum on Pedagogical approaches.....	38
4.4 The effects of curriculum on learning and social outcomes.....	44
General summary.....	46
Chapter 5: Assessment.....	47
Introduction.....	47
5.1 Key terms in assessment.....	47
5.2 Main issues in assessment.....	48
5.3 Research on assessment.....	49
5.3.1 Validity and Reliability.....	49
5.3.2 Assessment to support teaching and learning.....	51
5.4 Effects of assessment on pedagogical approaches.....	55
5.5 Effects of assessment on learning and social outcomes.....	60
General Summary.....	60
Chapter 6: Pedagogy.....	61
Introduction.....	61
6.1 Research on pedagogy.....	61
6.1.1 What makes a good teacher?.....	61
6.1.2 Teacher knowledge of the subject.....	67
6.1.3 Different groups in the classroom.....	71
6.1.4 Class size and organisation.....	74
6.2 The role of the teacher.....	78
6.3 Teacher professional development.....	84
6.4 Effects of pedagogy on students' learning and social outcomes.....	87
General summary.....	87
Chapter 7: Synthesis of Chapters 2, 3 and 4.....	89
Introduction.....	89
7.1 Learning and social outcomes.....	90
Chapter 8: Gaps and recommendations for Research.....	99
References.....	105
Appendix One: Literature sources.....	117
Appendix Two: Consultants and search protocol.....	119

Chapter 1: Overview

This review reports the findings of a literature search on the effects of curricula and assessment on pedagogical approaches and on education and social outcomes. This chapter outlines the structure of the review and comments on some important features which have influenced the search process and the subsequent analysis.

Chapter 2 describes the search process and outcomes. The literature search used a national and international team of experts, and was guided by a preliminary analysis of factors developed by consultation. The search used both electronic data bases and physical searches of key journals, and was iterative, with early articles providing direction for further sources. The criteria for selection of articles had a significant influence on this review. These were to be recently published (concentrating, though not exclusively, on the last five years) and major reviews or large empirical studies. These criteria focused this review largely away from New Zealand research, and smaller qualitative research which can enrich understanding of this complex area. In some cases where this was helpful the review refers to research not strictly within the criteria. The field is very complex and over 570 articles were processed for the review. This chapter describes how early in the process the complex nature of the interactions between curriculum, assessment and pedagogy was recognised as important, and the role of the teacher in students' learning and social outcomes seen to be central to the developing analysis.

Chapter 3 describes the development of an analytical framework for the review. This framework was developed by considering the nature of effective learning, and conditions which would enhance this. Effective learning is dependent on the role of the teacher in the classroom, influenced by curriculum, assessment and views of pedagogy. These influences were seen as sometimes being strong, and either working together or in conflict. The role of the learner and of influences outside the school are also considered as affecting the learning environment. At the end of Chapter 3 the analytical framework is presented (Figure 2).

Chapter 4 describes research on curriculum. The main considerations are whether curriculum should be mandated or open, what should be in the curriculum, and the appropriate relation between curriculum and assessment. Issues to do with curriculum content, the provision of different curricula for different groups, the development of high stakes subjects and activities are explored. The crucial and strong interaction between curriculum and assessment is outlined, with a focus on the influence of high stakes assessment procedures. The effects of curriculum on pedagogical approaches is discussed with reference to this interplay between assessment and the intentions of the curriculum. Consideration of the effects of curriculum on students' learning and social outcomes is deferred until Chapter 7 because of this complexity.

Chapter 5 describes research on assessment. The first section presents the research argument for significant learning gains from formative assessment procedures, and notes the substantial changes in classroom practice required for formative assessment to be introduced. The next section outlines research on the effects of assessment on pedagogical approaches, where features of the analytical framework are clarified. Consideration of the effects of assessment on students' learning and social outcomes is again deferred until Chapter 7 because of this interaction.

Chapter 6 describes research on pedagogy by first addressing the attributes of a good teacher. Shulman's analysis (Shulman 1988) of the knowledge bases for teaching is outlined and explored. The particular importance of teacher knowledge of their subject is discussed, and issues of classroom size and organisation explored. The central role of the teacher in a system of curriculum, assessment and pedagogy is then argued, leading to a discussion of teacher professional development. The effects of pedagogy on students' learning and social outcomes is briefly considered before the main discussion in the next chapter.

Chapter 7 brings together the review outcomes from Chapters 4, 5 and 6. Learning and social outcomes are described as resulting from combined influences from curriculum, assessment and pedagogy, occurring in the teaching and learning environment through the intervention of the teacher. This description uses the analytical framework for the review.

Chapter 8 considers gaps in our knowledge of the review focuses and makes recommendations for further research.

Chapter 2: The review process and outcomes

Introduction

This chapter first outlines the review requirements in Section 2.1. Section 2.2 documents the process of the review commenting on the complexity of the review task and noting some important features of the literature reviewed. At the end of this section the complexity of the interactions between curriculum, assessment and pedagogy, and the central role of the teacher in the review focus is briefly described.

2.1 The review requirements

The following aims and specific questions are from the review contract.

Aims of the review

1. To gain an understanding from the literature of the effects of the scope, content and nature of different kinds/levels of curriculum specification on:
 - pedagogical approaches and
 - education and social outcomes
2. To review the effect of different pedagogical approaches on education and social outcomes.
3. To identify and compare the effects of different kinds of assessment regimes reported in the literature on pedagogical approaches and education and social outcomes.

Specific Questions to Guide the Review

1. **The effect of curricula on pedagogical approaches and students' learning and social outcomes.**
 - According to the research, what are the most significant effects of different kinds of curriculum specification (for example, outcomes based) on :
 - pedagogical approaches; and on
 - students' learning and social outcomes, particularly those for Maori and Pacific students, students from low decile and immigrant backgrounds and students of differing abilities as well as the learning and social outcomes of boys as compared with girls?
 - What conclusions can be drawn from research which compares the effectiveness of pedagogical approaches in systems with a mandatory curriculum and those with other curricula arrangements?
 - What new light can be shed on established pedagogical beliefs and approaches in the classroom in relation to curricula policies and to learning and social outcomes?
2. **The effect of assessment regimes on pedagogical approaches and students' learning and social outcomes.**
 - What does the research say about the extent to which the presence or absence of particular assessment regimes influences students' learning and social outcomes?
 - What are the most significant effects of assessment regimes on pedagogical approaches reported in the literature?
 - What does the literature say about the effects of assessment regimes on teaching practice at different levels of schooling, for example, primary and secondary levels?
 - What does the literature say about the effect of different kinds of assessment regimes on the student groups identified above?

At the outset of the review the core project team noted that these research questions provided an apparently linear cause-effect structure, and did not acknowledge the interactions between assessment regimes and curriculum. The review process clarified the interactive nature of

curriculum, assessment and pedagogy, and the multifaceted effects of these interactions, mediated by the teacher and the learner, on learning and social outcomes.

2.2 The process of the review

A preliminary framework to examine the interrelationship between the variables was developed to assist in the search process. People with extensive experience in curriculum and assessment both nationally and internationally helped develop search templates which were used in conjunction with the appropriate search engines. The project team met frequently to audit the searches and collate and analyse the information.

The following **Variable Analysis of Factors** (Figure 1) was developed using the variables listed in the tender document under the appropriate headings of curriculum, assessment, pedagogy, educational and social outcomes. Under each of the headings variables that could be investigated were determined. Changes were incorporated after negotiations with the Ministry of Education team, findings from the literature searches, and key search directions given by the national and international search and guidance teams. The arrows on the diagram represented our understanding of the tender requirements with respect to the direction of the effect of the interrelationship between the variables to be analysed. **Keywords** (Table 1) were developed to assist the search.

Figure 1

Variable Analysis of Factors

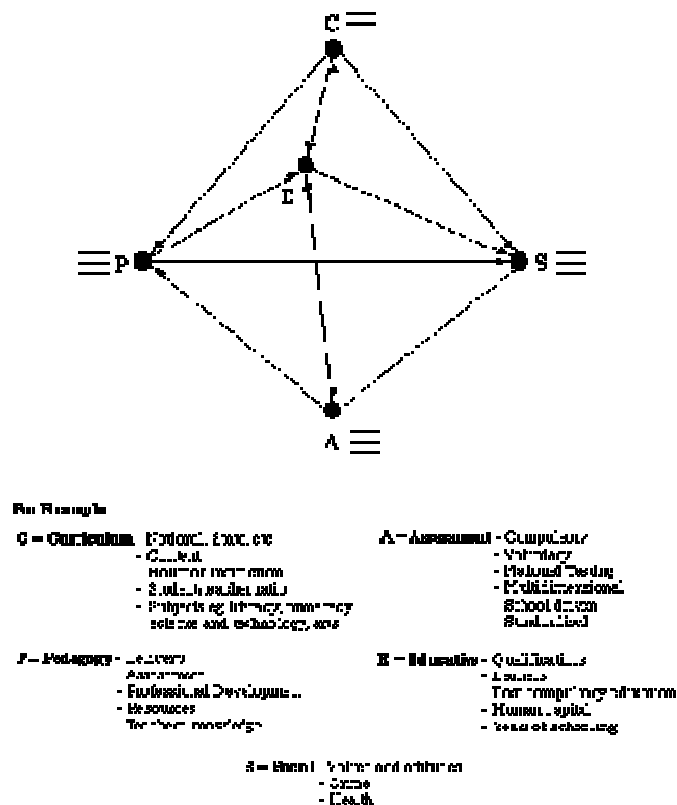


Table 1

THE KEYWORDS USED IN THE ANALYSIS
(primary keywords or **nodes** in the model are in bold)

Pedagogy		
delivery	Maori education	indigenous
first nations	multicultural	immersion
ESOL	assessment	formative
(second) language acquisition	summative	diagnostic
professional development	textbooks	resources
teachers knowledge-content	IT	education outcomes
teachers knowledge-pedagogical		
Educational outcomes		
Maori	indigenous	first nations
multicultural	immigrants	Maori students
Pacific Island students	gender	low decile students
disability	qualifications	barriers
years of schooling	human capital	curriculum outcomes
students of differing abilities	low achievers	
e.g. gifted		
Curriculum		
national	state	mandated
voluntary	national standards	state standards
school standards	hours of instruction	student teacher ratio
subject areas	outcomes	national outcomes
state outcomes	individual outcomes	curriculum aims
outcome based curriculum	knowledge based society	life long learners
education for a global society	multicultural	
Assessment		
assessment systems	formative assessment	summative assessment
national assessment	authentic assessment	testing
monitoring	evaluation	standardisation
compulsory	school based	qualifications
influence	teachers	teaching
pedagogy	learning	effects
values	different populations	attitudes
learning outcomes		
Social and Economic Outcomes		
self reliance	employment	parenting/family
human capital		

The process of the literature review

The literature review was iterative. Searching, analysing and writing were simultaneous over the period of the project with searching concentrated at the beginning. The process is described below.

Criteria for selection

The criteria used to select articles, books and reports for reviewing and summarising were:

- the presence of material addressing the links and relationships between the nodes in the preliminary framework (that is between curriculum, assessment, pedagogy, educational outcomes and social outcomes)
- major review articles within the nodes.
- meta-analyses of case studies within the nodes.
- articles by key researchers within the nodes.
- analytical commentaries on broad surveys such as the IEA and TIMSS studies.

Appendix 1 provides details of the sources for this review, listing the electronic data bases searched and the journals which have been physically searched on the recommendation of an overseas consultant, Paul Black, who noted the limitations of electronic searches arising from the varied definitions of the keywords used to organise the electronic data.

Initial search

The initial search used the information personnel as a scoping exercise to establish the nature and content of the literature. The overseas consultants were approached in person and by electronic mail. This provided valuable information through identification of key research and collection of recent unpublished material. There was overlap between this consultation and the developing draft research protocols.

Development of search protocol(s)

A draft search protocol was developed, commented on and refined by the core project team and the University of Waikato/National and International search and guidance teams in an interactive manner. This method enabled the project team to isolate the variables to be investigated. Key questions, open-ended questions, areas of investigation, key words (see Table 1 above), and questions were used to address the key international review literature.

Final protocol(s) to Waikato and international groups

The final protocol(s) were distributed among search and guidance teams. The protocols were returned and subsequently analysed. Literature known to the search and guidance teams was provided to the core project team for review and analysis. **Appendix 2** lists all the consultants and contains some information sent to members of the national and international teams. Background information on the project, a diagram of the preliminary framework being used in the project, a summary of the nodes and of secondary keywords related to the nodes, a protocol for response detailing information requested and a cover sheet to be filled in for each article was provided.

Searching for variables using the preliminary framework, the protocol(s) and library search engines.

The key words and structures were communicated to the information professionals for literature searching using the multi-dimensional search engines described in Appendix 1.

Review of initial outcomes

The identified abstracts from the searches were analysed and paper copies of appropriate literature sought.

Commentary on progress from National and International teams

Progress on the literature searches, and the subsequent key findings and trends were communicated with the search and guidance teams for feedback and further suggestions of areas of importance to be investigated.

Further refinement of search strategies

As trends and issues become apparent further searches were undertaken as appropriate.

Analysis and categorisation of findings

Analysis of findings was undertaken by the core project team.

Writing and review

The analyst/writer wrote the report in association with the core project team. The developing report was subjected to a review process. In the final week of preparation of this report a draft was provided for review by the core project team and Professor Noeline Alcorn, Professor Russell Bishop and Professor Ted Glynn, all of the University of Waikato School of Education. Their feedback was incorporated into the final report.

Comments on the literature search

The searching was comprehensive. Important features of the search process were the use of keywords developed from the contract preliminary framework to search electronic sources, following up key references suggested by the project team and national and international consultants, snowballing from these references, and physical searches of key publications. In addition the most recent IEA study was downloaded from the internet. Each of these processes provided important material. An internal check is that the same references have been located by several approaches, the value of a variety of search methods has been that each method has uncovered new and important references.

Due to a breakdown in the National Library system many interloan requests were delayed by several weeks. Some material which seemed very relevant to the review was requested from other countries but did not become available, for example, key research on the National Curriculum in England and Wales, and research on effective schooling from the United States. The review reflects the material which was obtained through the process described above.

The Waikato/National and International Search and Guidance teams

As indicated earlier, contact was made with all those listed on the national and international teams, and an initial response received from all members. The international team provided valuable information about key papers and researchers which both overlapped with the contract team's profiles and enriched them. A valuable commentary was provided by John Olson. Consultations have continued about key work to be considered, and refinement of the search protocol. The national teams have also contributed valuably to the information being processed providing confirmation of the literature already being analysed and enriching the data base.

2.3 Issues related to the complexity of the review process

This review covered a wide range of the recent literature, was confined to large scale studies, was inevitably biased due to some strong research foci, and did not find information about some important features of the New Zealand scene. The complexity of the interaction between curriculum, assessment and pedagogy, and the central role of the teacher became quickly apparent. These issues are briefly considered below.

The enormous compass of the task

The focus of this literature review is extremely wide. As a consequence the relevant literature is overwhelming. Material related to curriculum, assessment and pedagogy, and their effects on classroom learning and social outcomes is diverse and voluminous, and throughout the search and analysis new and significant material emerged. We completed an analysis of over 570 articles, books and reports identified in the searches, and 175 were analysed in depth for inclusion in the review. Due to the complexity of the interrelationship between the literature variables this process was more complicated than initially anticipated and required more input from the research team than the contract reflects.

This report should be viewed as an attempt to bring together a vast field of recent (last five years) literature in a short time frame. Further analyses of this body of literature, and perhaps further searches will be required to clarify many of the issues that are raised in this review. Even so, certain policy directions can be drawn from the gaps analysis and research directions outlined in Chapter 8.

Problems with reaching conclusions from the research

A common comment in the major reviews of sections of the literature relevant to this research is that overviews can not reach clear conclusions on the basis of the existing research data. The interplay between curriculum, assessment, pedagogy, learning and social outcomes is interactive, complex and difficult to analyse for policy advice. There are, however, some significant features of

this interplay, summarised in Chapter 7, which look to be particularly significant in understanding the education system and in considering improvements to it.

Problems with the research processes reported in the literature

Some of the research in the surveyed literature was empirically based used experimental or using quasi-experimental design: other research studies were based on self-report data, particularly about the effects of different curricula. In the self report studies, claims of educational improvement are not always supported by data from more precise measurement. Conclusions drawn in this review should be supplemented by empirically-based studies with case studies, preferably longitudinal, so as to gain insights into detail, links and relationships. This would provide the detail about individual actions and responses necessary to explore fully the many interacting variables. Chapter 8 elaborates this view.

Some strong biases in the articles reviewed

This field of research has been dominated by the response in the UK to their new National Curriculum and the associated assessment processes. This has been one of the most studied (and imitated) curriculum and assessment innovations. The inevitable consequence has been a flood of research documentation. This review therefore has a bias to commentary on the UK scene. This is not irrelevant to the aim of the review as New Zealand has been one of the imitators noted above. However, a good deal of literature from other parts of the world is included, particularly North America and Australia.

The research that applies to particular school subjects has a bias towards mathematics and science, which is not surprising since there has been a strong international research emphasis on these two subjects in last five years. The major international assessment regimes are the International Education Assessment (IEA) programme and the associated Third International Mathematics and Science Study (TIMSS). International comparisons have been strongly influenced by these probes of part of the school system, and much of the discussion about assessment has this bias.

The Eurocentric nature of the research

Our conclusions will be influenced by the often unquestioned Eurocentric nature of many of the major review and framework articles. For example, many of the review articles had neither an explicit cultural or class framework, nor reviewed research studies that dealt with specific groups of students. In addition, the research in classrooms did not often measure or report outcomes for the different groups present. We looked to supplement the main searches with specific searches for articles that look at the multicultural dimension in and between the nodes.

The complexity of interactions affecting learning

Many of the reviewed studies had a restricted focus because the field is so wide. We found substantial reviews of studies on each of the nodes of assessment, curriculum, pedagogy and on interactions within and across these nodes. Frequently these studies ended with pleas for more research, since important issues were still not clear after the analysis. This was due to the complexity of the many variables, and the imprecision in showing how they had been controlled. When considering the effects of these nodes on learning and learning outcomes the complexities are compounded. The large scale studies were often broad surveys of very disparate research in which variables related to classroom practice (such as teaching expertise, the teaching approaches used, the nature of interaction with students, the assessment processes and the consequences for individual students of all these factors) were invisible. The problem for the reviewers, and for our analysis, is that these and other factors appear from research to be crucially important in their effects on learning. Therefore the major surveys often reach tentative conclusions and warn that there is difficulty in associating a cause with the effects described. Case studies are seen as important to provide localised, contextualised knowledge to elaborate and explain the findings from large, quantitative surveys.

The central role of the teacher in a system

Few studies clarified the links as articulated in our search model. Where there were studies of the links we found that mediation by the teacher was crucial.

The field being surveyed is complex and interactive, systemic in the current description (see Olson's essay in Chapter 6, Section 6.2). This refers to the idea that change in one part of a system affects all parts of that system. Therefore, when educational reforms are developed, it is considered important to discover how different parts of the system will interact. Research on this view of reform, termed 'systemic research', considers the teacher and teacher development to be a significant factor since the teacher is at the centre of, and mediates the systemic change. Insights into appropriate interventions need to be gained through major longitudinal studies where the many variables can be explored through a combination of quantitative and qualitative data. Further insights might be gained through case studies to provide a rich data base. Previous researchers have commented in similar vein. Black and Wiliam (1998) comment in their major review of research on formative assessment and classroom learning:

It might seem desirable, and indeed might be anticipated as conventional, for a review of this type to attempt a meta-analysis of the quantitative studies that have been reported. The fact that this hardly seems possible prompts a reflection on this field of research. Several studies which are based on meta-analyses have provided useful material for this review. However, these have been focused on rather narrow aspects of formative work, for example the frequency of questioning. The value of their generalisations is also in question because key aspects of the various studies that they synthesise, for example the quality of the questions being provided at the different frequencies, is ignored because most of the researchers provide no evidence about these aspects.

Individual quantitative studies which look at formative assessment as a whole do exist, although the number with adequate and comparable quantitative rigour would be of the order of 20 at the most. However, whilst these are rigorous within their own framework and purposes, and whilst they show some coherence and reinforcement in relation to the learning gains associated with classroom assessment initiatives, the underlying differences between the studies are such that any amalgamations of their results would have little meaning.

At one level, these differences are obvious on casual inspection, because each study is associated with a particular pedagogy, with its attendant assumptions about learning; one that in many cases has been constructed as the main element of the innovation under study. There are however deeper differences; even where the research studies appear to be similar in the procedures involved, they differ in the nature of the data which may have been collected - or ignored. The fact that important determining features are often given no attention is one sign of the inadequate conceptualisation of the issues involved, indicating a need for further theory building. (p 53)

This review found much to reinforce Black and Wiliam's comments in relation to the foci of curriculum, assessment and pedagogy, and their effects on learning and social outcomes. Many of the reviewed studies explored different approaches and purposes, and were unclear in their conceptualisation of important variables (such as ethnicity of students, and the teaching and learning approaches used in the research study), and in their accounts of the interactions between the many interacting influences (such as clarifying the curriculum operating in the research study in relation to official curricula, and the role of the teacher in interpreting the many assessment processes). For these reasons, particularly the systemic nature of the review focus and the central role played by the teacher, the review team sought a framework to structure the analysis. The factors contributing to this framework will be outlined in the next chapter, which concludes by describing the framework for analysis.

Chapter 3: Towards a framework for analysis

Introduction

Development of a framework for analysis first focused on the teaching and learning process, since this is where the interaction between curriculum, assessment and pedagogy occurs, and where the learning and social outcomes are promoted. Section 3.1 focuses first on the New Zealand Curriculum Framework, arguing that achievement of the many aims of this Framework requires optimum conditions for effective learning. The main features of a good teaching and learning process are then developed by reference to the current literature about learning experiences, the role of the teacher and the influence of assessment procedures. Section 3.2 provides further comment on the role of the teacher as interacting between curriculum, assessment and the classroom. Section 3.3 addresses the role of the learner. Section 3.4 briefly considers influences from outside the classroom on students' learning and social outcomes. Section 3.5 sets out and explains the framework for analysis.

3.1 Effective learning

The contract seeks information to inform New Zealand educational policy regarding the influences on learning. The New Zealand Curriculum Framework consists of two parallel documents in Maori and English and describes seven learning areas: Language and languages (Te Reo Maori), Mathematics (Pangarau), Science (Putaiiao), Technology (Hangarau), Social sciences (Tikanga-a-iwi), the Arts (Nga Toi), and Health and physical well-being (Hauora). They also include essential skills (nga tino pukenga) of: communication, numeracy, information, problem-solving, self management and competition, social and co-operation, physical, and work and study. In addition a statement is provided about attitudes and values (nga waiaro me nga uara) to be addressed when learning. The New Zealand curriculum is the formal knowledge structure that New Zealand teachers start from and integrate into school organisational structures. Given the two statements it follows that curriculum, assessment and pedagogical research involves the inclusion of different knowledge and language.

Among the English language curriculum documents there is considerable similarity between these aims and those for many other countries about expected learning outcomes from schooling (the research literature on curriculum is explored in Chapter 4). What is not included are the issues and learning outcomes surrounding bi-culturalism and bi-lingualism.

The first focus of the review is to describe the features of classrooms which would provide optimum conditions for good learning outcomes. There is substantial international and national consensus about some of these features (for example see, Adey, 1997; Duncan and Welch, 1996; Gipps and MacGilchrist, 1999; Harlen and James, 1996; Ladson-Billings, 1995; Newmann and Wehlage, 1998; New Zealand Learning in Science (LISP) and Learning in Technology Education (LITE) projects).

This review makes the assumption that an important aim of education and schooling is to bring about learning with understanding (Harlen and James, 1996). This has been called learning as an interpretative process or deep learning. Therefore, learning with understanding is actively understood and internalised by the learner and makes sense in terms of a learner's experience of the world. Learning is not simply a collection of isolated facts which have been memorised. It differs from rote learning in that it is linked to previous experience and can be transferred to new situations. What is known and understood will, of course, change with new experience and as new ideas and skills are presented to help make sense of it. Harlen and James (1996) characterises learning as:

- progressively developed in terms of big ideas, skills for living and learning, attitudes and values;

- constructed on the basis of previous ideas and skills;
- applied to contexts other than those in which it was learned;
- owned by the learner in the sense that it becomes a fundamental part of the way he or she understands the world;
- simply not ephemeral knowledge that may be memorised for recall in examinations and subsequently forgotten.

There is a further dimension to learning which is highlighted by the disciplined inquiry approach in the UK reviewed by Gipps and MacGilchrist (1999), who note that things are easier to learn if they make sense. Meaning makes learning easier because the learner knows where to put things in their mental framework, and meaning makes knowledge useful because likely purposes and applications are already part of the understanding. This is incorporated in socio-cultural views of learning where learning is a process of knowledge construction. Teaching is not seen as a direct transfer of knowledge but as an intervention in an ongoing knowledge building process. It is recognised that much of learning is of a social nature and focuses on an individual acquiring knowledge-in-social-action. Therefore social views of learning see knowledge as an entity, something cohesive and holistic which then provides a scaffolding for later learning.

This view of learning recognises that features such as gender, class and culture help to shape our society. Women, ethnic or racial groups, and the poor are seen as located unequally in the social formation and education systems. The knowledge which they offer and the practices which constitute them, are seen as being unequally incorporated into teaching and learning. Teaching and learning includes not only subject knowledge and practices, but recognition of broader social relationships. From this point of view, it is argued that any concept taught must be concerned with what is taught, how it is taught and how it is learned and, more broadly, with the nature of knowledge and learning (Luke & Gore, 1992; McCarthy & Crichlow, 1993). While pedagogy is often seen as the components of the act of classroom teaching, socio-cultural views of learning see it differently. To Gipps and MacGilchrist (1999), 'knowledge-in-social-action' can be seen as the knowledge that is produced, negotiated, transformed and realised in the interaction between the teacher, the learner and the knowledge itself.

To promote effective learning, Harlen and James (1996) argues that what is needed are learning experiences that:

- are well matched to the existing point of development of the ideas, skills, attitudes and values;
- have continuity with, and build on, previous experience;
- relate to current interests and experience;
- are perceived by learners as relevant, important, stimulating and valued for themselves, rather than simply for their usefulness in passing tests and examinations.

Other researchers would suggest that these characteristics fall short in their conceptualisation. Ladson-Billings (1995) and others go further to suggest that learning experiences should be embedded in the social relationships of the wider society.

There are bodies of knowledge that societies expect teachers to teach and students to learn. There are, of course, continual debates about *what* knowledge and *whose* knowledge is included in any curricula. The curriculum has been targeted by different groups as being a place of contestation, for example, the inclusion of Maori knowledge, drugs programmes, and ICT initiatives. However, the exponential increase in the amount of information, coupled with the rapid change in the nature of employment, indicate that there should be far greater emphasis on learning which can be transferred and applied to new situations. This requires us to look at strategies related to learning.

From his review of relevant research, Crooks concludes that 'there seems to be a strong case for encouraging the development of deep strategies from the early years of the educational system'

(Crooks, 1988 as cited in Harlen and James, 1996, p 447). Crooks refers here to the simple but powerful way of identifying the approaches to learning which lead, on the one hand to understanding and on the other to rote memorisation, in terms of the distinction between *deep* learning and *surface* learning. Deep learning has the following characteristics: an intention to develop personal understanding; active interaction with the content, particularly in relating new ideas to previous knowledge and experience; linking ideas together using integrating principles; relating evidence to conclusions. Surface learning has the following characteristics: an intention to be able to reproduce content as required; passive acceptance of ideas and information; lack of recognition of guiding principles or patterns; focusing learning on assessment requirements. Between surface learning and deep learning, Marton et al (1984, as cited in Harlen and James, 1996) also proposed an intervening category which they called 'strategic learning'. This reminds us that efficient learning is often a combination of both surface and deep learning, for if we were to learn everything in depth we would have time to learn very little. Likewise, if everything was surface learning we could hardly describe ourselves as educated at all.

The provision of effective learning experiences with these kinds of characteristics, as defined by Harlen and James (1996), depends on the teacher:

- having a thorough and deep understanding themselves of the subject matter to be taught, how students are likely to learn it and the difficulties and misunderstandings they are likely to encounter;
- having a clear idea of the progression in the ideas, skills etc. which are the goals of learning and the course students are likely to take in this development;
- being able to recognise the point in this development reached by their students;
- knowing and being able to use various strategies to find out and to develop students' ideas skills, etc.

Teachers cannot provide experiences and activities that guide students' progress towards understanding of ideas if they themselves do not know what these ideas are; nor can they provide experiences that challenge students' misunderstandings if they themselves share the same misunderstandings (Harlen and James, 1996). Although teachers generally have sound pedagogic skills to carry them through 'difficult' aspects of the subjects they teach, the results for students are restricted and impoverished learning opportunities, particularly for learning aimed at understanding. For example, recent research into teaching science at the primary level indicated that the coping strategies adopted by teachers with poor background knowledge included relying heavily on prescriptive workcards or texts giving students step-by-step instructions, emphasising process aims and neglecting conceptual development, minimising students' opportunities to ask questions, avoiding practical work and any equipment likely to go wrong (Harlen, Holroyd and Byrne, 1995).

The presence of different cultural groups in the classroom raises the question of what constitutes culturally relevant teaching. The international literature considers two issues associated with culturally relevant teaching. Firstly, whether students from diverse cultural backgrounds interact differently with the teacher, with other students and with the subject matter under study. Secondly, whether teachers' expectations of students significantly determine performance. Our search found few investigations of how different pedagogy effects change to individuals or different groups in the classroom. New Zealand is rare in having introduced language immersion education in the language of the involuntary minority (Ogbu, 1995). Our search found little information about the consequences of immersion education here or internationally.

Some key aspects of pedagogy are elaborated in Chapter 6. These include: teachers' knowledge in and about the subject they are teaching; teachers' knowledge about the cultural perspectives of the learners in their classrooms; teachers' expectation of their students; and where teachers have low expectations of children, they provide a more restricted learning opportunity.

Assessment is an integral part of the learning process and has both formative and summative functions. These two sets of functions are mainly a matter of when they occur in relation to their purpose, and not a differentiation of rigour or quality. Formative assessment is an on-going informed interaction between the teacher and student designed to enhance student learning. Therefore it provides feedback to the teacher and to the student about present understanding and skill development in order to determine a way forward. To use information about present achievements in this way means that the progression in ideas and skills must be in the teacher's mind - and as far as possible in the students' - so that the next appropriate steps can be considered. Formative assessment is discussed in detail in Chapter 5. Summative assessment describes learning achieved at a certain time for the purposes of reporting to parents, other teachers, and the students themselves. Summative assessment is discussed later in the review. When curricula have been developed in other languages appropriate assessment procedures need also to be developed. Assessment of Maori immersion programmes occurs through direct translation of assessment items for the English curriculum. Crombie and Johnson (1998) question the match between Maori language immersion classroom practice and translated assessments.

The teacher and learner are central to learning in the classroom. The next two sections briefly discuss the related roles.

3.2 The intervening role of the teacher

Curriculum, assessment and pedagogy interact in a complex manner and the teaching and learning process of the classroom is at the focus of strong and frequently contradictory forces. Olson argues (Chapter 6, Section 6.2) that the role of the teacher in change is a central issue for understanding the way curriculum, pedagogy and outcomes interact, and the process of educational reform.

The teacher is crucially at the centre of this system because the links between;

- curriculum and pedagogy,
- curriculum and assessment,
- pedagogy and assessment,
- curriculum and learning outcomes,
- curriculum and social outcomes,
- pedagogy and learning outcomes,

come from the teacher playing a major role in interpreting curriculum, in fashioning pedagogy, and in devising assessment activities. In all of these roles the teacher reflects his or her past history, and implicit and explicit theories of teaching and learning. The conflicts, risks and contradictions involved in any reform to the system necessitate what Anderson (1998) describes as the 'collaborative reconstruction of a new social ground' for teachers. Teacher professional development (see detail in Chapter 6, Section 6.3) is an important factor in these effects of the teacher on the teaching and learning environment.

This raises another area that seems to be missing in the research field, namely a focus on teacher change within a socio-cultural framework. Whilst there is a large amount of literature looking at culturally appropriate strategies that may take account of the student learning, the role of the teacher has been left largely untouched with the exception of trying to learn new strategies. For example, educators are often advised to become familiar with various cultural practices and/or cultural knowledge so that they might intervene appropriately in a pedagogical manner. This suggests that any classroom encounter can be simply "managed" as a pedagogical moment requiring racial, cultural or gender sensitivity, which further implies that communication between the participants, teacher and student, is simply technical. Furthermore, it is thought that misunderstandings in the classroom arise because the parties are culturally, racially, physically, mentally, or sexually different. All educators need to do is to learn how to work their way through these differences. Two USA research studies provide evidence of the importance of teacher expectations in addition to technical matters. Christine Sleeter (1993) has argued that in the USA the adoption of apparently helpful cross-cultural strategies does little to change white teachers'

views of their Asian and African-American students as being capable of the same level of achievement and range of desires as their white students. Second, in a study of excellent teachers of African-American students, Ladson-Billings (1995) found that the teachers shared a common philosophical and ideological underpinning of their practice. These included how they thought about themselves as teachers; how they thought about others, including their students, the students' parents, and other community members; how they structured social relations within and outside of the classroom; and how they conceived of knowledge.

3.3 The role of the student

In the same way as the teacher is crucially central to the system the student has a significant effect on the teaching and learning process. This occurs through the influence of knowledge and attitudes brought to the classroom, and through active interpretation of and interaction with classroom activities. These influences may support or impede learning. Students have developed habits (dispositions) which may be hard to change, and which may interfere strongly with intended reforms.

Nuthall (1997) provides an overview of student thinking and learning in the classroom, based on New Zealand and international research. He gives examples of classroom interaction from the following New Zealand studies: Alton-Lee and Nuthall, (1992), Nuthall and Alton-Lee (1992) and Alton-Lee, Nuthall, and Patrick, (1993). The main points in these reviews are:

- while it can be assumed that there is a link between teaching and learning, it is not easy to quantify learning gains and identify their source;
- learning is mediated by students. The way that tasks are structured, the questions that teachers ask, and the examples that students practice can only have an indirect effect on student learning;
- teaching needs to be viewed as the management of the classroom as a community, rather than simply as instruction in specific curriculum areas;
- variables that affect student thinking about subject matter in the classroom are,
 - how classroom tasks and practical activities engage and structure cognitive processing,
 - how social processes of the classroom influence learning and knowledge construction, and
 - the role of the teacher in motivating, structuring and guiding student cognition;
- learning is embedded in the cultural nature of classroom processes.

This overview sees language use in classrooms as working to structure the social relationships between students and teachers and to create the knowledge and meanings that constitute the experienced curriculum. Curriculum knowledge and the conditions for thinking about experience and acquiring knowledge are created in the process of using language. Their commentary is summarised as follows: 'All the evidence we have now about how language shapes classroom experience, about how social processes structure the content of what is talked about, and how students remember both the context and content of their classroom experiences, all these make it clear that it is the whole of what goes on in classrooms that determines how students think and learn'.

Other studies complement these findings and provide further insights into the role of the student in learning. Gipps and Tunstall (1998) report a study of 6 schools in 5 LEAs in London, involving 8 teachers of year 1 and year 2 children (infant children) across these 6 schools. Forty-nine children were selected from the eight classes for detailed study. In each class, two high attaining, 2 low attaining and 2 average pupils were chosen for interviewing and observation. The study ascribes significant influences on learning to the learner's: self-esteem, attributions for success and failure, motivation, feedback received from others e.g. peer responses, formative assessment, self-evaluation of academic achievement, self-image, self-concept, perceived failure, and perceived success. For example, in a related study Tunstall and Gipps (1996) looked at students' perception of teacher feedback. Their major research questions about classrooms in which formative

assessment was being used to support learning were: what sort of feedback do teachers give children; and how do children interpret, understand and act on this feedback? The results demonstrated that students understood a wide range of teacher feedback in infant classrooms, and could articulate freely shared, evaluative experiences and self-monitoring strategies. In this case the students appeared to have attitudes consistent with the intended changes to classroom practice.

3.4 Influences from outside the school

The research questions for this review are posed in such a manner that influences on students' learning and social outcomes are to be analysed as consequences of curriculum, assessment and pedagogy, in other words from events which predominantly occur in classrooms and learning institutions. Comment on the influences on students' learning and social outcomes from outside the classroom are not required. Another literature review has been sought on "The impact of family and community resources on student outcomes".

The literature surveyed made many references to factors such as family background and support, socio-economic effects, and cultural attitudes to learning which can significantly affect an individual student's learning. These are not assembled into an argument here as this is being prepared elsewhere, but the review team would concur with an analysis which finds strong interactions between family and community resources and student outcomes. Teachers in classrooms can make a difference, and can develop appropriate pedagogical approaches to counter students' difficulties arising from any negative impacts of family and community resources. They are, however, often insufficiently resourced and supported to overcome some severe consequences of factors outside their immediate compass.

3.5 A framework for the review analysis

The previous discussion in this chapter has focused on the features of good learning and a good learning environment, and described the interplay between curriculum, assessment and pedagogy which affects the environment in which the process of teaching and learning occurs. This interplay influences the process of teaching and learning, and the teacher has a central role in the provision of a good learning environment, influenced by the student as discussed in Section 3.3.

The next four chapters provide reviews of the literature on curriculum (Chapter 4), assessment (Chapter 5), pedagogy (Chapter 6) and the combined effects of curriculum, assessment and pedagogy on students' learning and social outcomes (Chapter 7). These reviews were shaped by the framework developed for the review (Figure 2). In this framework curriculum (which is developed in response to the many views of stakeholders) is seen as influencing teaching and learning through its specification of learning goals, its definition of learning outcomes, and through the variety of targets (importantly ethnic, social and gender differences) addressed or ignored in curriculum definition. Assessment is seen as contributing to teaching and learning through two processes, its focus on learning and on evaluation of the education system. Assessment is seen as influenced by factors within and outside the school. Stakeholders are represented as influencing the assessment processes. Pedagogy is seen in this framework as having major interactions with teacher knowledges, and as being dependent on teacher professional development. Pedagogical approaches for different classroom groups need to be analysed. On the left-hand side of this framework broad arrows link the three key features of curriculum, assessment and pedagogy, and indicate that interactions between them can be enabling or disabling, together or in conflict. This feature of the framework provides scope for analysis of how, for example:

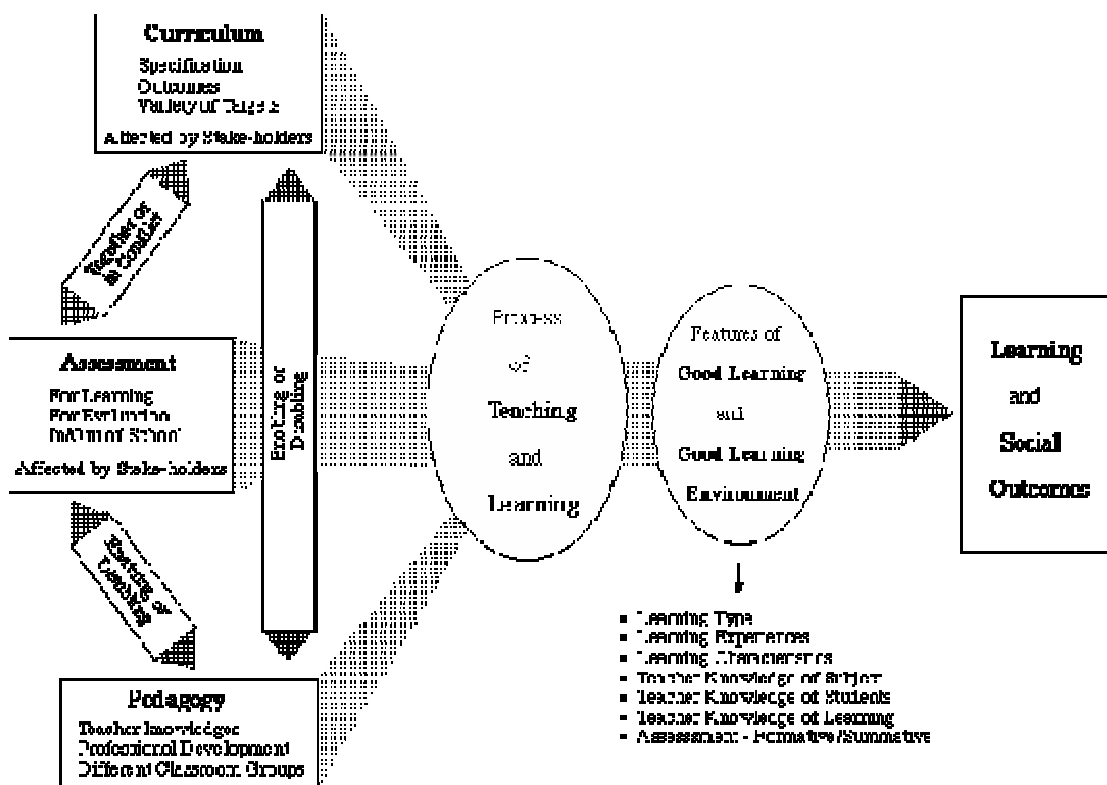
- assessment procedures may conflict with curriculum intentions through causing an excessive emphasis on some aspects of the overall curriculum.
- curriculum statements may enable or disable student-centred pedagogical approaches.
- assessment procedures may enable or disable particular pedagogical approaches.

The process of teaching and learning is depicted in the framework as dependent (importantly as a result of decisions made by the teacher) on the interactions between curriculum, assessment and pedagogical approaches. Learning and social outcomes stem from this process of teaching and learning, and the analysis was given some direction by seeking from the literature features of good learning and good learning environments (related to the features listed below the small arrow in the framework). The remainder of this review will make frequent reference to this framework in order to bring together the substantial body of literature surveyed.

The framework, then, assisted the review team to clarify the review requirements. The next four chapters flesh out this framework.

Figure 2

An analytical framework for the review



Chapter 4: Curriculum

Introduction

This chapter begins with organising statements about the varieties of curriculum specification and the many curricula which result from interpretation of official curricula, and then discusses the stakeholder influence on curriculum development (Section 4.1). Section 4.2 details research on curriculum through consideration of three questions:

Should curriculum be mandated or open?

What should be in the curriculum?

What is the appropriate relationship between curriculum and assessment?

In considering the second question the focus is on curriculum content; the provision of different curricula for different groups in the school; the pressures of new subjects and the definition of required further conceptual knowledge and skills; the relationship between achievement and the curriculum; and the relative importance of subjects and the status of different features of the curriculum. The third question raises the crucial strong interaction between curriculum and assessment. Section 4.3 details research on the effects of curriculum on pedagogical approaches. Section 4.4 indicates that consideration of the effects of curriculum on students' learning and social outcomes will, because of the systemic interactions described in the analytical framework be deferred to Chapter 7.

4.1 Commentary on curriculum

In examining curriculum internationally, and how curriculum, assessment and pedagogy interact, we needed to be aware not only of different types of curriculum specification but also that there are different levels of interpreted curricula. There are different levels of curriculum specification at the state, region and school level across countries but although there are analyses of different types of curriculum specification at the country/state level (for example the research by NFER), there appears to be very little comparative research evidence of the influence of curriculum specification on student achievement.

Official curricula are at the start of a cascade of interpreted curricula. These interpretations begin when supporting material (of all kinds) is devised to provide an official curriculum. Other stages include interpretation of the curriculum by teacher organisations, by school systems, by individual schools, by subject teachers, by the classroom teacher, and finally the students' interpretation of the curriculum as a consequence of teaching and learning activities. This complexity of responses to a curriculum, and the many change influences throughout the cascade mean that it is difficult to find direct and robust links between official curricula and the end result in terms of learning and social outcomes. Put in another way, the intended curriculum becomes an implemented curriculum and then an achieved curriculum, and considerable modification is possible in this chain of events. The systemic nature of interactions between curriculum, assessment and pedagogy further compound the situation and make research and commentary on curriculum a difficult process. Hence this literature review found that there is little research which would provide robust answers to questions about the nature of curriculum, the appropriate mix of knowledge and skills which curricula should contain, and the impact of curriculum on learning.

The social process involved in devising curriculum

New Zealand has a Curriculum Framework and individual curricula which have been devised over many years. Any national or official curriculum reflects the particular needs of the society for which it is devised. These documents are a complex response to the perceived educational needs of this society, contributed to by groups with differing and sometimes conflicting views of appropriate content. The process is political involving contestation and debate. It is not surprising then that while research played some part in this process it is important to recognise that the

compromises reached in devising the New Zealand curriculum were not the result of rigorous application of research [see for example Bell, Jones and Carr (1995) on the development of the English language version of the New Zealand science curriculum, and McKinley (1995) for its subsequent impact on the Maori version of the science curriculum].

4.2 Research on Curriculum

This literature search found few research investigations of curriculum, with more investigations being related to the interplay between curriculum and assessment, and between curriculum and pedagogy. Nevertheless, valuable reports about curriculum and related matters by Paul Black and his colleagues (in major surveys of OECD studies and of the international testing associated with the International Education Assessment - IEA, Third International Mathematics and Science Study - TIMSS) provide a framework for this commentary.

Crucial issues for curriculum policy could be posed by three questions:

- Should curriculum be mandated or open?
- What should be in the curriculum?
- What is the appropriate relationship between curriculum and assessment?

Research which provided commentaries and partial answers to the first two questions resulted from this literature search; some important contributions to understanding about the third question will be discussed.

4.2.1 Should curriculum be mandated or open?

Most countries now have detailed and mandated curricula, national or statewide, whereas a few have a more classroom-based response to general guidelines. The most relevant information for this review is contained in the review by Black and Atkin (Black and Atkin 1997) of TIMSS data (related to science and mathematics curriculum and learning outcomes) for the 13 OECD countries (Australia, Austria, Canada, France, Germany, Ireland, Japan, the Netherlands, Norway, the United Kingdom [Scotland], Spain, Switzerland, and the United States) which were also involved in their curriculum research (Black and Atkin, 1996). Their main findings can be summarised as follows.

The potential educational advantages of national standards and a national curriculum, are not clarified by the TIMSS and the OECD study, but it should be noted that it was not one of their goals. However, three major messages emerge.

- There is some form of national science standards in many countries, but by no means in all of them.
- National standards, where they exist, differ from country to country on at least two dimensions: the concepts they identify for students to learn and the actual influence they exert on teachers and students.
- No country is satisfied with the condition of its existing programs of science education, even those that have adopted national standards and that score near the top in comparisons of educational achievement.

Black and Atkin found no strong evidence from the TIMSS data that the existence or absence of nationally mandated curricula led to improved performance. While accepting that comparisons are difficult, they note that the median rank order in science scores at eighth grade is 17 for countries whose curricula are not centralised, 19 for those regionally centralised, and 22 for those nationally centralised. These figures reflected the fact that, although eight of the top 10 countries had national curricula, so did eight of the bottom 10. In mathematics again, eight of the top 10 and nine of the bottom 10 had national curricula. They note that curriculum is associated with the historical and cultural tradition of a country, and the question of whether a change in a particular country would produce a benefit cannot be answered by comparing the status quo across different countries. On this point about mandated curriculum one of the authors states in this review that he

had had close experience of the imposition of a national curriculum in England and Wales. This author comments that there was little evidence that this approach has yielded improvement and much evidence that it had damaged the status and morale of those on whom improvement most depends, the teachers.

The authors note also that there are no official statements of national subject-matter standards in some countries, including Germany, Switzerland, and Canada. Here curriculum definition occurs almost entirely with cantons, states and provinces – sometimes with individual schools and teachers. In Spain there are guidelines, but only for those schools under direct control of the national ministry. Such schools, however, are not the norm because education is commonly a responsibility of each of the autonomous regions.

There does not, then, appear to be a clear answer to this question of whether curriculum should be mandated or open, to guide policy. These comments are relevant to other sections of this review when the curriculum is considered as a conceptual framework which could clarify the work of teachers (Section 4.2), and when the effect of assessment on pedagogical approaches is considered in Chapter 5, Section 5.4.

4.2.2 What should be in the curriculum?

This question can be addressed through topics such as

- how subject curriculum should be framed (including the provision of differing curricula for different groups in the school system),
- the relative importance of subjects,
- new subjects in the curriculum,
- the status of different features of a curriculum.

Some research relevant to each of these themes is discussed in the following sections.

4.2.2.1 Subject curriculum (including the provision of differing curricula for different groups in the school system),

Notwithstanding our earlier comment that curriculum development is largely a political process, curricula as frameworks to guide learning would be expected to be influenced by research. Harlen (1999) surveyed curriculum research in science for the Scottish Council for Research in Education, and noted that the research (which has mostly considered conceptual development) provides the following conclusions about science curriculum planning (which may have some general relevance):

- Attention should be given to the complexity of the reasoning which is demanded for understanding particular phenomena when deciding where these should be placed within the curriculum framework.
- There is evidence that some ideas (common across science) are more easily developed in some content areas than others and their generalisation should not be assumed.
- Research evidence about the conceptual development of pupils should be used in planning and sequencing curriculum guidelines.

The research base implied in Harlen's summary is very recent and in some areas its interpretation remains disputed.

Is student achievement related to curriculum specification?

Another approach to curriculum development could be to survey international curricula and examine trends. So when deciding on curriculum content an elegant solution may appear to be adoption and adaptation of subject curricula in successful countries in international tests - putting aside for the moment the relevance of rankings made on such crude bases (see Chapter 5). The research of Black and Atkin (1997) warns against any simplistic international comparisons stating

that 'just as there is no magic bullet for improving an education system, so there is no clear path to be found by trying to imitate 'successful' neighbours.'

This conclusion is again based on the TIMSS which measured student achievement in 45 countries and also analysed curricula, instructional practices and certain teacher characteristics; and the OECD study in 13 countries [Australia, Austria, Canada, France, Germany, Ireland, Japan, the Netherlands, Norway, the United Kingdom (Scotland), Spain, Switzerland, and the United States].

While the TIMSS centred exclusively on the current status of education in the subjects in question the OECD study focused entirely on innovations.

Significant points for this review are:

- In the 13 countries that participated in the OECD study there is universal unhappiness with the present state of science and mathematics education. All countries (no matter where they were placed in terms of their TIMSS ranking) are trying to make changes to their existing programmes. Even apart from the OECD countries, countries such as Japan (which scored highly in TIMSS) viewed their students as being uncreative and were seeking reforms to encourage problem-solving and first-hand original investigations¹. Although there are differences in the reasons for innovations, virtually everywhere the curriculum is becoming more practical and related to everyday life.
- All the countries in the OECD study were found to be revising their standards toward applications, toward practical work, toward increasing students' capacity to see real-world relevance, and toward enhancing students' enthusiasm for further study of the subject. The authors note that there is little in the TIMSS tests to measure such outcomes, and that if these goals were important then it became harder to know which countries to imitate.
- The move seen in the OECD study toward greater emphasis on the practical was checked against the TIMSS students' reports on the frequency of practical work in their classrooms. On this index, Japan is ranked about sixth, the US 11th, Germany 21st, and Hungary 39th. The OECD study also showed a marked effort to enhance emphasis on real-world problems. TIMSS provided two indicators here. One was the importance teachers give to the understanding of how science is used in the real world, where Hungary was fifth, the US 11th, Germany 15th, and Japan 33rd. The other indicator was the students' reports on the frequency of using things from everyday life, where again the US was well ahead of Germany and Japan, but Hungary dropped to 32nd, between Germany and Japan.
- Another important issue centres on the views of students. For students' reports about liking the sciences, the US response produced 71% liking or liking a lot, whereas Hungary was at 62% and Japan and Germany at 56%. In order to achieve high scores on the particular types of test used for TIMSS, US students might have to be taught in such a way that they would like the subject far less than they did at present. The authors conclude that given the diversity of the indicators, no clear direction for achieving high scores could be inferred by any simple analysis.

Black and Atkin make important points about the discussion reported above. They comment that it is all too easy for a discussion such as theirs to be driven by a model of international competitiveness. It was notable that, in the OECD studies, international comparisons of student

¹ In this context Fujioka and Suwannaprasert (1995) surveyed 600 teachers, university lecturers, mathematicians, scientists, engineers on what they would like in the mathematics curriculum. In general the existing elementary mathematics curriculum emphasises knowledge and intellectual abilities and skills to the exclusion of general skills, attitudes, interests and appreciation. The sample surveyed desired changes in the Japanese mathematics curriculum towards improved attitudes, interest in and appreciation of mathematics.

scores were only one of the 'demons driving reform' – and then in just a few countries. Their general conclusion was that there is no clear link between success on test scores and some of the other goals for improvement, and no clear indication that the top-scoring countries were outstanding in meeting other criteria that many people consider important. They provide two reasons for caution about this analysis. First, it may be that complex multivariate analyses of the data will indicate some clearer correlations between performance scores and other indicators. Even if this were the case, any complex correlation would open a logical gap to be crossed only by a qualitative judgment as to causality – complicated profoundly by the leaps across cultural contexts. Second, in the drive for excellence, the TIMSS scores could only be part of the picture. If the cost of high scores was to exacerbate weaknesses on other important criteria, then difficult choices would have to be made. Tradeoff costs might be unavoidable. Nevertheless international comparisons using IEA studies do not provide clear guidance on how curricula specification might impact on student achievement. Furthermore, a curriculum specification – achievement link needs to take account of the very complex matter of what happens when teachers implement the specifications, for the variation is, potentially, enormous.

The politics of curriculum development can sometimes be divisive. Fore (1998) reports how, after more than a year of debate, the Virginia State Board of Education adopted new curriculum 'standards' in June 1995 for elementary and secondary school social studies. The debate over the content and shape of these standards was acrimonious. At odds were two communities with fundamentally different conceptions of the nature of knowledge and the purposes of social studies: on one side were the state's conservative Republican Governor and his political appointees to the State Board of Education, and on the other was the professional social studies community, comprising college educators, curriculum supervisors and some classroom teachers. The Governor and his team won this debate by managing both the discourse and also the official structure in ways that largely excluded the professional community from substantive participation in the standards-development process. The authors describe this episode as a case study in the politics of exclusion.

Continuing the focus introduced above, the review surfaced a complex discussion about education for 'Effective Citizenship' and its incorporation into the curriculum. These arguments are outlined below as they introduce issues to do with integrated curriculum content or separate subjects, and with teaching in the area which provide a case study relevant to other curriculum innovations.

Effective citizenship as a case study

Education for effective citizenship is broadly located in social studies, and concerns the education of citizens to take their place in society. The issues relating to citizenship education have general application and the ideas they suggest can be applied to different aspects of the curriculum.

The first issue is should citizenship be taught as a separate subject or should it be part of the general school programme? The Task Force on Civic Education (1998), Print (1996), The Australian Civics Experts Group (1994) and organisations such as Civitas argue for a separate subject with a sequential programme. Topics in the US National Standards for Civics and Government (1994) statement include : What is government and what should it do? What are the basic values and principles of American democracy? How does the government established by the constitution embody the purpose, values and principles of American democracy? What is the relationship of the United States to other nations and to world affairs? What are the roles of citizens in American democracy?

Proponents of a separate subject approach have a good deal of influence, but the alternative view of integration is also strongly promoted. The England and Wales Qualifications and Curriculum Authority (1998) argues that while citizenship education is an essential component of education, it should be taught in combination with other subjects rather than separately. Lohrey (1996) takes this argument further claiming that citizenship is a generic skill rather than a body of content knowledge. Cogan (1997) also argues that effective citizenship is best taught through the general ethos of the school.

Mellor and Elliot (1996) claim that pedagogy is 'the most important factor in teaching and learning effective citizenship'. On the issue of pedagogy there is also disagreement between a proscribed pedagogy and an open approach. Formal programmes of citizenship such as those suggested by Civitas and the Civics Experts Group suggest an agreed pedagogy. However most commentators argue like Fearnley-Sander and Sprout (1996) for a variety of teaching styles. Hughes (1994) claims that there is no best model of pedagogy for teaching citizenship and Kennedy (1996) claims on the basis of a major review of international literature, that inclusive programmes which are student centred and incorporate valuing are the most effective.

The other major issue in the area of citizenship education is how to reconcile a socialising and essentially conservative civics programme aimed at conformity with the increasing diversity of communities. Garbutcheon Singh (1996), Kennedy (1996) and Parry (1998) point out the need for Australian schools to rethink the nature of citizenship, and to redefine it so that it caters for the ideas and aspirations of different cultures and minority groups. Other arguments come from the United States, particularly Gay (1997), Kaltsounis (1997) and Parker, (1997). While these arguments are not based on substantial research, these writers have acknowledged expertise in this field and considerable influence on curriculum design in the United States and internationally.

Complementary to this case study is the major report by Whitty, Aggleton and Rowe (1996) which explored curriculum provision in social education. As discussed above this curriculum focus can be introduced through a theme base or a subject base. The important point they make is that a subject base for curriculum usually has an evaluation system based on more formal evaluation techniques such as tests and examinations than the less formal (and more suited to abstract concepts) methods of assessment for themes. An important consequence is that the subject organised curriculum carries higher status in the school system, through being seen as more academic.

New subjects and the crowded curriculum

Internationally decisions about curriculum have produced shifts in definition of subjects. A frequent innovation is a separate curriculum for technology and/or information studies, other subjects have been integrated - environmental studies- or integrated studies have been separated. The content of knowledge and skills within subjects has also been modified. This further compounds international comparisons, and within countries makes comparisons between present schooling and schooling in the past (when curriculum demand was less) fraught with analytical difficulties, except for those who assume rose-coloured spectacles. The many pressures on schools can be argued to result in an overcrowded curriculum.

In the New Zealand context, the work of teachers and the consequences for children's learning need to be set against what teachers are required to teach in the available time. Primary teachers are typically responsible for one class of students for most of the hours of schooling in a year, and all the subjects or curriculum areas. The maximum time available in a week is about 22 hours and there are seven curriculum areas to be taught - about three hours for each. Subjects like English and Mathematics usually receive more time.

Curriculum areas are made up of several strands which are often 'sub subjects'. For example, the English curriculum statement has three strands, each with two major components: oral language, written language, and visual language. Inside these strands are yet more complex clusters of content and skills to be mastered, such as learning to spell and handwrite and use correct grammatical forms. The science curriculum statement divides content into six major strands, requiring a teacher to have, ideally, a wide range of knowledge in various aspects of science. Maori knowledge has been introduced into the English versions of the curricula, which further extends the teacher knowledge required in teacher training and in private study.

Altogether the seven areas and their attendant strands represent a much greater array of content than in the past, with the school year being the same. In addition there are many extra demands on

the available teaching-learning time, such as responding to social needs and increased outside school activities.

New Zealand has also introduced curriculum statements which define new subjects - Technology- and reorganise other content - Health and Physical Well-being. Computing has also made an impact on schools. These responses to perceived social needs have not been researched in terms of their overall impact on the school system. Certainly schooling now is very different from schooling in previous decades, and the consequences of this pressure is not well researched.

Conceptual knowledge and skills

The New Zealand Curriculum Framework defines essential learning areas, essential skills, and attitudes and values. In Chapter 5 on assessment the issues associated with achievement of knowledge and skills, and the encouragement of attitudes and values is considered. In this context of curriculum it is worth noting a problem about the development of many of these attributes of learning (such as self-management and competitive skills, social and co-operative skills, and work and study skills) identified in the curriculum. These skills, seen as important by the pressure groups contributing to curriculum development, lack validation through research support. Adey (1997) provides rare evidence for critical thinking skills when reporting significant achievement gains when there is an appropriate balance between the subject area and the development of critical thinking. Assessment measuring achievement in these attributes of learning is a complex process which lacks a research base. This research gap means that curriculum change is not supported by evidence. An example of the possible problem is that curricula have provided statements of progression in learning which are based on little in the way of research evidence.

Black, Brown, Simon, and Blondel (1996) in a major review looked at underlying progressions in pupils understanding as they move through a particular domain of learning. In two areas of science and mathematics, pupils' difficulties appeared to arise from a lack of understanding of certain key concepts. The authors argue that teachers should be directed to these concepts. The authors found considerable differences among children of the same age in terms of their understanding of concepts, for example some seven year olds were at least as advanced as many thirteen or fourteen year olds. These findings have major implications for both curriculum design and teaching. The authors argue that the National Curriculum for England and Wales does not appear to significantly organise the concepts into levels of difficulties and this affected what the children learnt. Revising the conceptual targets would mean that teacher assessment would have to be re-focused. However the authors point out that it may not be easy to design written items to assess conceptual steps in a reliable way.

In an extensive research project spanning seven years, Case, (1991) and colleagues investigated the nature of children's conceptual structures in relation to mathematical thought, social and emotional development and spatial development. The research was based on the neo-Piagetian theory that children develop conceptual structures within different domains, but that there are simultaneous general systematic changes which more broadly affect the children's abilities to employ these capacities in real life situations. Drawing conclusions from the wide variety of experimental studies that make up this project, Case (1991) concludes that although intellectual functioning exists within certain domains, they are linked by commonalities. These commonalities are representative of central conceptual structures and it is through these structures that the forces of biology and culture meet and children develop an understanding of the world around them (Case, 1991).

Overall this study explores the conceptual framework of children's thought and knowledge, but it does not explicitly link this with curriculum at any point. On the other hand, the results of the individual studies, such as the section on the development of children's logico-mathematical thought (including development of scientific reasoning and acquisition of everyday mathematical knowledge), could easily be applied to the development of a curriculum.

The area of progression in all subjects needs more research to clarify the issues raised here.

Different groups in the school system

The most striking feature of New Zealand curriculum documents from an international perspective is the provision of parallel documents in Maori and the system of schools and sections of schools promoting learning in Maori. Few systems have responded to different cultural groups in their society in this way, and research informing this issue is rare. The issue of Maori achievement is reported in Chapple, Richards and Walker (1997) where the lack of research is a significant finding. Research related to cultural issues in the classroom in other systems is now discussed.

Many countries are still debating the inclusion of different cultures and their knowledges into the curriculum. Snively and Corsiglia (1998) consider science curricula, and argue that the definition of science is a de facto "gate keeping" device for what can be included in a school science curriculum. They argue that in most science classrooms around the world, western modern science has been taught at the expense of indigenous knowledge, and that the intense philosophical debate around the inclusion of multi-cultural science in mainstream science education has not always been sufficiently wedded to practical science knowledge issues. All too frequently educators have developed science programmes and curricula with little understanding of embedded cultural assumptions, and the philosophical debate that has taken place has occurred in relation to few concrete examples. They describe the burgeoning traditional ecological knowledge (TEK) and indigenous knowledge (IK) literature which now provides educators with conceptual stepping-stones to rich collections of multi-cultural science perspectives, approaches, examples and data. These should assist educators to design science curricula that explore critical science, technology, and community issues related to sustainability and bio-regionalism.

These authors argue for the inclusion of TEK in school based science education programs to provide examples of local and relativist, multicultural science. Learning science is viewed as culture acquisition that requires all students to cross a "cultural border" from their everyday world into the subculture of science. They provide a conceptualisation of science education that attempts to take full account of the multi-dimensional cultural world of the learner. The article contains numerous examples of "time proven, productive, and cost effective indigenous science", and specific recommendations for helping *all* (their italics) students move back and forth between the culture of Western modern science and the culture of local indigenous science.

In the USA Allen (1997) examined indigenous models for integration of science and culture in curriculum and instruction. The study was conducted over an eighteen month period during which the researcher worked with 13 Native American Indian Schools in curriculum reform. In the summer of 1996, the Four Directions Challenge in Technology Project brought together teams of teachers, administrators, community members, and students from 13 Native American Indian schools for a 2 week institute in culture, technology, and curriculum development. The teams produced thematic curricula and multimedia projects that incorporated science and culture. Classroom discussions, electronic journals, informal interviews, and curriculum products were used to determine group priorities and concerns, models for cultural integration, and effective ways to support local reform. The research report comments that although teams often shared areas of concern - and thus targeted similar science content - the models used for curriculum development differed according to community values concerning culture and instruction. The report suggests providing communities with continued instruction in curriculum design and encouraging local control of content.

A feature of the search for material on culture and the curriculum was that only one research report was found which considered provision for gifted students in different cultural groups from the main stream. Friesen (1996) argues that giftedness in First Nations culture (Blackfoot were the focus for this project) has different characteristics than those currently recognised in education. Characteristics of giftedness would here include: humbleness, aural/oral memory, non-competitive, concrete learning situations, traditional ways of dealing with personal issues (e.g. singing), and family, religious or spiritual activities taking precedence over school functions. The emphasis is on

Indian heritage, culture and language. Main stream practice is stated not to have recognised these characteristics so that there is an under-representation of 'giftedness' in gifted programmes. Friesen notes that for a Zia Pueblo woman the characteristics of giftedness would include the cultural knowledge of food preparation, language, bread-making and sewing, and sharing of her time, knowledge and efforts.

This discussion of the research findings on the curriculum is the appropriate section to explore another targeted group in the classroom; the boys. Curriculum implications have often been drawn from considerations of their achievement.

'Underachievement' of Boys

Curriculum change is often argued as a response to the perceived needs of different groups. For example the New Zealand science curriculum has a major introductory section on the participation and strategies to use with particular groups of students. The under-achievement of boys debate is included here to illustrate that the response may need to be complex, and that simplistic strategies may not be appropriate.

The release of analyses of assessment results in many countries typically produces public comment about the performance of males and females, frequently this is highly-coloured and seeks changes to the curriculum. The New Zealand Ministry of Education commissioned the report *Gender Differences in achievement and participation in the compulsory school sector* (Ministry of Education, 1999) to look at the available data. The Ministry Report states that male achievement in education had not declined over the reviewed period (1986-1998). This report recognises that gender differentials exist, but describes these as consistent over this time.

Key findings of the report were that:

- male students make up 75% of all school suspensions.
- the trends for retention rates beyond school leaving age have been similar over the 1986-1998 period, with male retention rates slightly less than female.
- the number of students leaving school with no qualifications has decreased overall and, in 1997, this stood at 19.8% of males and 15.3% of females.
- the participation rates in School Certificate (SC) is very similar.
- slightly more Year 12 females participate in Six Form Certificate (SFC), although the participation rates for both sexes are above 74%.
- a marked gender differentiation still occurs in the ten most popular School Certificate subjects for females and males. This follows stereotypical gender lines, for example, workshop technology and technical drawing occur in the male top ten while typing and home economics occur in the female top ten subjects.
- females outperform males on most indices related to literacy and language. However, male performance has not declined over the time reviewed as the differences in achievement have consistently favoured females.
- at school leaving age males outperform females in mathematics literacy. This gender difference is also found in School Certificate Mathematics, and Bursary Mathematics with Statistics, but not Bursary Mathematics with Calculus. The performance of males in mathematics has not declined over time.
- males outperform females on science literacy at school leaving age. Approximately equal proportions of males and females have attained top marks in School Certificate Science, while in the sixth form females outperform males in every science. Bursary subjects show Biology and Physics slightly favour females, while Chemistry slightly favours males in achievement.

The report concludes that any differences in participation and achievement are not new and that male performance has not declined over time. Of concern are levels of suspension, participation and achievement in English for males.

The causes of under-achievement

Debate about gender and achievement is not confined to New Zealand [examples are Reed (1999) for England; Yates (1997), and Gilbert & Gilbert (1998), for the Australian context]. Of interest are the descriptive characteristics of the ‘underachieving boy’ and of suggested ‘solutions’ through changes to curriculum and pedagogy. This discussion parallels the multicultural debate; the two discourses intersect through seeing boys as a differing cultural group.

Any ‘crisis’ in relation to boys and schooling is defined by performance in assessment tasks during the time of compulsory schooling. Boys’ achievement is posed in relation to girls’ achievement, with claims for a relatively stronger performance of girls in many areas of the curriculum. The greatest concern is about boys reading and writing. This debate using broad-based statistics on boys and girls ignores the intersection of social class and ethnic groups. Differential outcomes can result from differences in social and curricular experiences.

Measures being taken or considered to improve boys performance frequently address the concern about their leaving qualifications and social behaviours. More boys are leaving school with no qualifications, and boys feature predominantly in truancy, in school suspensions or expulsions, and in special education needs units and remedial reading classes. The contribution to these statistics of being Maori or from a low socio-economic group, or both or neither, is not known.

Explanations for any differences come from psychological and sociological disciplines, and parallel some multicultural arguments discussed in Chapter 3, Section 3.1. Reed (1999) notes that in England a popular explanation is that boys and girls are different by nature, bringing different aptitudes and orientations to learning environments. In this scenario it is claimed that:

- boys’ development is slower than girls.
- their mental processes develop along different paths, resulting in different structures and operations by adulthood.
- the male brain finds it difficult to deal with ‘reflective emotional-centred’ tasks and shows a preference for ‘speculative thinking and action’.

In other words, gendered preference is embedded in human development. The implication is that the teacher should identify individual pupil learning needs and match them to a preferred learning style.

Sociological explanations in this discourse are often simplistic and adhere to normative theories of social roles. Any change to curricula whether at the national or school level in response to different groups should not be based on simplistic analysis of normative theories of social roles.

4.2.2.2 The relative importance of subjects

Curriculum frameworks provide statements about what should be covered in schools; they are often less specific about the appropriate allocation of time and resources for the separate subjects and the various skill developments implied. There is some literature which explores this issue and notes the emergence of ‘high stakes curriculum’ subjects, frequently through responses to assessment procedures. Madaus (1988) describes how testing influences the curriculum (a fuller description is in Chapter 5, Section 5.4). He notes that in every setting where a high-stakes test operates, a tradition of post exams develops which eventually *de facto* defines the curriculum, and that teachers pay attention to the form of questions in high-stakes tests and adjust their instruction accordingly. A consequence can be that some subjects are perceived to be more important than others since more attention is focused on them in the testing regime.

The impact of assessment on the relative importance of subjects is complex, not well researched, and beyond the main scope of this review. Pollard, Broadfoot, Croll, Osborn, and Abbott (1994) cite classroom observations in English primary schools which found that, as a result of the new

national curriculum and associated testing processes, the core curriculum, particularly English and mathematics had become dominant at the expense of other subjects.

Plewis and Veltman (1996) provide more detail, and suggest that the other subjects were losing time whilst the core remained much the same. They looked at the time spent on different subjects in Year 2 (7 year old) classrooms in 22 schools, comparing the results with data previously collected in the schools. They found that time spent on working inside classrooms decreased compared with time outside the classrooms, such as in assemblies and physical education. One reason was that some teachers were given non-contact time for record keeping and assessment. Distribution of time across subjects showed writing, mathematics and reading made up 44% of classroom time, and 19% of school time - very similar to the mid-1980s. Much less time was spent on art, craft and construction activities; down from 20% in the 1980s to 10% in 1993. More time was being spent on science. In 1993 more time was spent giving lesson instructions, and there was less procedural-administrative time taken up. Therefore, the classrooms days become somewhat more "academic." However, there was no reduction in outside activities such as outside play and meal time.

In the same context Brown (1996) records that national testing as used in the England and Wales is high stakes, and found significant moves from integrating subjects towards subject based teaching, more regular formal testing of assessed material, and curriculum content being modified to match the testing. How this affected the distribution of focus between subjects is implied but not specifically explored.

4.2.2.3 New subjects in the curriculum

The context explored above is relevant to the introduction of new subjects into the curriculum. New Zealand has introduced two new learning areas, Technology, and Health and Physical Well-being into the New Zealand Curriculum Framework, with some elements of each coming from previous curricula. Appropriate questions would be whether these subjects have become high or low stakes, and what were the consequences for learning in them. The search found no information about these issues, which deserve attention if the expectations of curriculum developers are to be met by practice.

4.2.2.4 The status of different features of a curriculum

Black and Atkin (1997) report a general trend in the OECD countries in terms of increasing practical aspects of the subjects. Other innovations have included a shift to student centred learning, individual responsibility, using software and multi-media, social issues in science, and moves to integrate curriculum.

This diversification of curriculum raises issues to do with the status of features of a complex curriculum. Black and Atkin (1997) note that there are problems associated with: teacher subject knowledge in the new features; using assessment to enhance teaching; devising practical real world problems to support innovations; enhancing the problem solving capability of students; the closer relationship of school and community required in many innovations. In this context high stakes assessment processes (see the next section) can mean that content knowledge has a dominant status, and innovations are not implemented as intended.

4.2.3 The relationship between curriculum and assessment

The development of a curriculum would ideally encompass the simultaneous development of an assessment programme which fulfilled the intentions of the curriculum. In this way the curriculum and the assessment processes would enhance each other. The experience in many countries has been that this complementarity has been lost, particularly when a major function for assessment procedures has been to measure teacher and school performance. Then conflict can arise between

assessment to enhance learning and assessment for measuring achievement. It would appear therefore that assessment regimes have a powerful effect on school curricula.

The major analytical paper of Madaus (1988) on the influence of testing on the curriculum defined high stakes tests as those whose results are perceived by students, teachers, administrators, parents/care-givers or the general public as being used to make important decisions that immediately and directly affect them. High stakes tests can be norm- or criterion-referenced, internal or external in origin (which means that school assessments which may seem to be low stake can become high stake if they become enmeshed in important decisions about students and teachers).

Madaus lists six principles of the power of testing to influence the curriculum.

- 1) If students, teachers or administrators believe that the results of an examination are important, it matters little whether this is true or not - the effect is produced by what individuals perceive to be the case. High stakes assessment is all too often a symbolic solution to real educational problems. As test scores rise over time - as they must do if teachers teach to the test, then policy makers can point to the wisdom of their action and the public's confidence is restored.
- 2) The more a quantitative social indicator is used for social decision-making the more likely it will be to distort and corrupt the social process it is intended to monitor.
- 3) If important decisions are presumed to be related to test results, then teachers will teach to the test. High stakes tests can focus instruction giving students and teachers specific goals to attain. If the test is measuring basic skills, preparing students for the skills measured by the test could, proponents argue, serve as a powerful lever to improve basic skills. Unfortunately the only evidence to support this position is that the scores on tests of basic skills rise, not that the basic skills rise, nor that the skill necessarily improves.
- 4) In every setting where a high-stakes test operates, a tradition of post exams develops, which eventually de facto defines the curriculum.
- 5) Teachers pay attention to the form of the question on a high-stakes test and adjust their instruction accordingly.
- 6) When test results are the sole or even partial arbiter of future educational or life choices, society tends to treat test results as the major goal of schooling.

Madaus comments that the impact of high stakes tests on teachers

leads to cramming; narrows the curriculum; concentrates attention on those skills most amenable to testing; constrains the creativity and spontaneity of teachers and students and finally demeans the professional judgment of teachers (p. 85)

and that students can be affected since high stakes tests can have a negative impact on poor and minority students, resulting in increased drop-out due to failure in tests, and decrease in self esteem, responsibility, and assertiveness. This report cautions that careful study is required if high-stakes assessment is to be introduced as cautions are not particularly clear. Nevertheless, the Madaus findings have been echoed many times.

Black and Atkin (1996) acknowledge the same problem when, after they have noted the value to education of formative assessment, they discuss national assessment processes through summative assessment. They point out that a final function of summative assessment, which must be acknowledged, is to serve the right of all citizens to know how well their schools are performing. As always, there is a paradox. It is fruitless for the public - parents, politicians, journalists or whoever - to make contradictory demands. They must not expect schools to respond to new

opportunities and changing social needs while at the same time demanding that they perform to unchanging criteria and unchanged practices. The public needs to be educated to understand curriculum change, new forms of assessment, the link between the two, and that new opportunities for learning require matching changes in formative and in summative assessment. Their view is that citizens have a right to know that changes in their students' education are producing verifiable improvement. At the same time the teachers who carry the burden of innovation as they strive towards new and often unfamiliar aims have a balancing right to support from the society that employs them.

This interaction between curriculum and assessment is a recurring theme in the literature reviewed for this report. The review questions are concerned with the effects of curriculum on pedagogy; it should be emphasised here that these effects are tied closely to assessment and that this often drives the interaction between curriculum and the classroom.

Black and Atkin (1997) comment that external assessments in particular, whether statewide or national, have traditionally exerted powerful control over what teachers feel obliged to do. Yet the impact of such assessments on efforts to engage students in original and complex work can be devastating if the tests measure only memory. Then teachers' efforts to do more than teach to the test are discouraged, and public support for reforms becomes more difficult to generate.

They note that even tests that measure such thought processes as the formation of hypotheses, the design of investigations, and assessing the correspondence between conclusions and evidence, are not very helpful if they are given at the end of a course and do not provide feedback to the teacher while teaching is in progress. In this context they note the benefits in some participating OECD countries from using tests as guides throughout the instructional process. This trend, they see, growing from recognition that in cases where concept development is highly sequential, as in some mathematics courses, many students lose their way and have little hope of learning anything of consequence unless their basic difficulties are diagnosed and addressed promptly.

These authors report two significant interactions between curriculum and assessment implementation. In France, large-scale assessments are now being used for diagnostic purposes in mathematics. The tests are given at the beginning rather than at the end of the school year so that teachers know at the outset what their students must learn. The assessments are also intended to serve as models for teachers as they prepare their own tests.

Norway offers another new approach to testing. In that country, teachers involve their students systematically in self assessment. As self-evaluators, students must develop a far deeper grasp of what their learning is supposed to achieve. They must understand the aims of the learning and the criteria for success. Self-assessment of this sort meshes closely with other efforts to help students take greater responsibility for their own education.

Curriculum and assessment are therefore strongly interactive, either in an enabling or disabling mode, as represented in the analytical framework for the review (Figure 2). Mehrens (1997) in his review on the relationship between assessment and curriculum notes that high stakes assessments impact on both curriculum and instruction, but alone are not likely to be as effective as teacher development.

This topic is explored further in the next section where the focus shifts to more detailed interactions between curriculum and pedagogy.

Summary of 4.2

This section has explored the complexity of curriculum processes. No firm research-based conclusions can be given as to whether curriculum should be mandated or placed more in the hands of teachers. Curriculum is interpreted at different levels and therefore it becomes difficult to understand the relationship between curriculum and student achievement in the absence of detail about what happens in the classroom. A further complexity arises since curriculum reform often

occurs at the same time as assessment or other reforms. This makes it difficult to discover any causal relationship between variables of curriculum and student achievement.

This literature search found little research evidence to provide robust answers to questions about the nature of the curriculum, the appropriate mix of knowledge and skills that curricula should contain and the impact of curriculum on learning. All OECD countries are revising their standards towards more practical work, student relevance and motivation to learn. Even countries that scored well in TIMMS were revising their curriculum to include more problem solving, creativity, etc.

There is some form of national standards in most countries but little research to understand the impact of these. No country is satisfied with its curricula, even those who achieve well in international IEA type studies. There is no clear evidence that the existence or absence of nationally mandated curricula led to improved performance. However, there are other levels of curriculum specification at regional and school level and there is little or no research evidence to understand the influence of this. We know very little about what occurs in classrooms as a result of particular curriculum structures. This is a significant research gap internationally.

Evidence is provided that indicates that when critical thinking skills and content are included in a school curriculum student achievement can be raised in formal examinations.

The area of progression is one of the least understood areas, yet this is crucial for designing curriculum and enhancing student achievement.

Curricula are often influenced by the type of assessment regimes put in place and it becomes difficult to isolate the variables contributing to student achievement. High stakes assessment can drive the curriculum content, the relative time spent on subjects and the pedagogical approaches. This is discussed in detail in the next section.

The definition of subjects and skills in the curriculum and the interplay of new subjects with existing subjects and approaches has been explored. The interpretation of official curricula into classroom practice is strongly affected by assessment processes as indicated in the analytical framework for this review (Figure 2). The important relationship between curriculum and assessment is the focus of the next section.

4.3 The effects of Curriculum on Pedagogical approaches

The literature reviewed provided detail of the interplay between curriculum and assessment, and their impact on pedagogy. This feature of the analytical framework for the review (Figure 2) is now described.

Doyle (1992) explores the link between curriculum and pedagogy in a meta-analysis of several hundred studies. Historically, pedagogy and curriculum have been researched as separate entities. Curriculum research has tended to centre on theoretical issues such as the difference between disciplines and the knowledge that should be included in the curriculum. On the other hand, research on pedagogy grew out of the behavioural psychology discipline and thus has focussed on teaching behaviours. Largely missing from this field of research is any recognition that while teachers are affected by the dictates of the official curriculum, classroom action represents the curriculum experienced by students (Doyle, 1992). Doyle (1992) describes the growing component of research on tasks and task structures as providing useful insights into the way curriculum is translated into action because tasks or activities are the basic building blocks of classroom action. However, he notes that the way teachers interpret the curriculum needs to be explored in much greater depth because curricular reform efforts continue to be based on the assumption that changes to the official curriculum immediately translate to changes in classroom activity. In reality, curriculum prescriptions have less of an effect than supposed, because classrooms have their own inherent structures composed of interactions between teachers and students (Doyle, 1992).

Pollard et al (1994) reports findings from the Primary Assessment Curriculum and Experience (PACE) project which investigated the effects of the then new National Curriculum on English Primary schools. This longitudinal analysis of 54 children as they moved through 9 schools (from 1989/90 to 1995/6) entailed classroom observations, teacher and pupil interviews, teacher questionnaires and head teacher interviews. The paper describes the broad impact of change in terms of values, understanding, and power and presents an analytical framework covering dimensions of change. Preliminary findings revealed several contradictory findings - change and resistance, commitment and demoralisation, and decreased autonomy and professional development.

The introduction of the National Curriculum was seen as seriously compromised as it alienated many teachers. The reforms introduced new constraints and regulations into almost every area of teachers' work. English teachers had been gradually evolving a view of knowledge and learning which they considered allowed them to emphasise their own professional skills, judgment and understanding without undue external influence. The introduction of the new curriculum exerted pressure on schools to introduce more subject specialist teaching. This was considered inappropriate by many teachers, as was the pressure to treat assessments as providing reliable categorical evidence. Teachers' responses are reported as varying between compliance, mediation and resistance. The research records increasing stress, a declining sense of personal fulfillment, and pessimism about the future amongst many teachers who felt burdened by what they saw as unnecessary imposition.

Classroom observations showed that the core curriculum, particularly English and mathematics had become dominant at the expense of other subjects. There was also a strong move away from combining subjects in classroom tasks, to a focus on single subjects. Teachers in the study increasingly directed pupil activities and shifted to more whole class teaching. (These shifts were against the pedagogical approaches previously developed in response to perceived pupils' needs). The first Standard Assessment Tasks (SATs) were found to require a great deal of classroom time and school resources, and to distort normal classroom practice. They were found to be difficult to interpret and implemented in variable ways, and teachers were concerned at the anxiety shown by pupils. In these ways, then, curriculum design impacted upon pedagogy, but whether they improved student achievement is uncertain.

Pollard et al (1994) report some more positive responses. The teachers were found to be broadly agreed about the structural benefits of having a National Curriculum to provide coherence and an organisational structure for schooling. Collegiality in schools was seen to have increased through cooperation about curriculum planning. Cooperation was evident during other responses to the many demands of the reforms (including ways to subvert the intentions of national assessment). Again, this sheds no light in student achievement.

Evidence has come however, from an analysis of later developments in the primary school by Gipps (1996). In this research the research team worked with 8 schools, chosen to represent typical UK schools. The key findings at key stage 1 (7 years) were that the national assessment resulted in:

- increased levels of discussion and collegiality in the 8 schools in the early stages of the implementation. Helping each other negotiate meanings for the SATs (standard assessment tasks) and in their moderation acted as a training device and broke down group barriers.
- raised professionalism of the teachers involved in leading assessment training and policy development. The teachers moved from being technicians operating in an imposed system, to being more knowledgeable on assessment.
- teachers spent more teaching time on the requirements of the SATs, e.g. punctuation, spelling, mental arithmetic. The experience of the teachers in the first year of implementation had helped them teach in a more focused way in later years.
- the teachers taught areas of the curriculum so that their children could do assessments on those topics. Therefore, the teachers taught to the curriculum as well as to the tests. A

rolling programme of attainment targets included in the testing meant that teachers could not become too narrowly focused in preparing pupils for tests.

- greater care in planning, close observation of children, and a more detailed understanding of individual progress was reported by over half of the 8 principals in the study.

These changes were ascribed by Gipps et al (1996) mostly to the performance type activities in the SATs, which have since been abandoned in favour of pen and paper tests.

In the same research at age 11 the impact on pedagogy of assessment and league tables for schools included:

- a change from mixed ability teaching to some form of streaming.
- a move away from cross-curricular to work towards more subject based teaching.
- a move to more whole class teaching.
- introduction of regular formal testing.
- a greater focus on reading, spelling and mathematics, in a way that they were likely to meet in the testing.
- preparation for the tests, through revision and looking at past tests.
- many changes to teaching style. Half the teachers in the sample of 29, advocated a mixed teaching style (transmission, discovery, constructing) but in the light of the testing 25% felt pushed into more transmission because of the amount of work to be covered by the curriculum or the tests. The teachers also felt unable to use children's ideas as starting points for the work and regretted this.
- the teachers also recorded positive effects related to: improved planning, design of teacher tasks, teacher knowledge of individuals, assessment procedures; and developments in basic teaching skills.

In summary, there was more teaching to the test than in Key stage 1 due to the high stakes of testing (league tables for schools). In contrast, because of the high stakes, the assessment had more impact on the teachers organisational and teaching practice. The researchers commentary on learning questions whether the higher levels reached in reading, spelling, and mathematics were a result of teachers spending more teaching time on requirements of the SATs, such as punctuation, spelling, and mental arithmetic. The SATs emphasised active, multimode assessment and interaction between teacher and pupil and provided a better opportunity for many students to show best performance.

These researchers note that some teachers felt they had lost control of the curriculum, though others said they stuck to their own beliefs and practice in assessing children.

Preece and Skinner (1999) discuss summative and formative processes in relation to the National Curriculum in England and Wales. They see a role for each type in informing teaching and the public, but note that if the two forms of assessment are to complement each other satisfactorily, they must have equal status and that their complementary roles be understood. An evaluation in 1996 reported that the delicate balance between the two types of assessment was in danger of being tipped in favour of the national tests and summative processes. These were found to be having a strong effect on classroom practice, and teaching to the test could undermine the validity of test results. It was therefore recommended that the testing programme be reduced by half, with one test per subject so that the tests could become a part of classroom life, without the intensive preparation for them evident at present.

Galton, Hargreaves, Comber, Wall, and Tell (1999) also studied the impact of national curriculum on junior school classrooms in England. Their results are in broad agreement with those of Gipps. Teachers in primary schools were found to do more talking and the pupils more listening than 20 years previously. By far the largest amount of this talk consisted of teachers making statements. Questions asked of the children were mainly recall of facts, or to solve a problem for which teachers expected a correct answer. Open, speculative or challenging questions were comparatively rare. In science, where the highest percentage of open questions was recorded, teachers were three

times more likely to require a single correct answer than they were to invite speculation. The demands of programmes of study in the national curriculum subjects appeared to require teachers to cut down pupil participation in order to get through the curriculum content.

This report argues that teachers used to feel respected. Now they thought of their work as just a job, where they were told what to do rather than using professional judgment. These authors also note a tendency in the UK to move whole class teaching where previously group or individual strategy might have been used.

Galton, Hargreaves and Comber (1998) found that in small rural primary schools the changes to pedagogy had not been as marked, and the teachers views indicated that they felt able to cope with the new curriculum and its knowledge demands. The authors speculate that this is due to these schools operating in clusters.

Alexander, Wilcocks and Nelson (1995) also found considerable changes in England and Wales classroom practice towards more curriculum planning and management, and in assessment and record-taking, but argue that this took place against a backdrop of relative continuity in pedagogy at the deep level, that is that teachers continued to teach with the same pedagogical approaches.

Whitty, Power and Halpin (1998) argue that progressive primary school teaching and schooling, certainly in England and Wales has given way to more traditional teaching and schooling. The government-defined processes of evaluation were seen to be regulating performance and controlling what was to be learnt. The authors argue that there has been an increasing emphasis upon what they call old-fashioned education and an increasing trend in England and Wales towards streaming of children.

Silvernail (1996) reports more consideration of the UK reforms. In interviews of 50 teachers, head teachers, educational officials and policy researchers in 1993 the majority of teachers were found to have fundamentally changed the way they teach, what they taught, and how they assessed their students. The amount of instructional time was directly related to the national testing that was to occur. The national curriculum was found to dilute the classroom curriculum by limiting the time available to explore new areas. The prescribed assessment programme was also found to contribute to narrowing the curriculum - thus the assessed curriculum is becoming the taught curriculum. With the introduction of national testing and league tables, temporary and permanent exclusions from classrooms had increased by 66%.

In an attempt to gain an understanding of the tensions and debates surrounding current patterns of curriculum reform, Prawat (1995) asked outstanding university based experts and teachers to critique current textbooks in their subject area and to respond to a series of questions concerning ideal curricular practices during indepth interviews. Three teachers and three researchers in each of the six subject domains (mathematics, science, social studies, literature, music and art) participated in the study. Although there was widespread agreement about the need to reform curricular practices so that students become actively involved in their own learning, the teachers tended to favour changes that favoured planning with the child in mind, while the university researchers tended to take more of a disciplinary approach focussing on issues such as content and instructional models. Prawat (1995) points to this difference of opinion about curriculum as a significant impediment to implementing curricular reform.

Clarke (1996) also looked at the impact of national assessment. In a review of three other studies, high stakes traditional tests were found to result in narrowing the curriculum, and encouraging teaching to the test. Replies to 400 structured questionnaires sent to year 2, year 6 teachers and head teachers were analysed. During the testing period the main changes in pedagogy were

- more silent working,
- a reduction in hearing children read and
- a reduction in teacher intervention and support.

Another effect was holding children back on earlier levels for testing rather than moving them onto new work. The tests are described as under-representing the national curriculum in reading, focusing on a narrow aspect of reading.

In this study 75% of teachers described increased distress in their classrooms and general anxiety and 50% of the teachers described distress at the level of the whole class. In relation to a focus on external tests, one third of the teachers prepared for the tests, another third said they would spend time revising tests from previous years. At key stage two teachers said they would spend more time revising and all teachers had prepared for the tests. The core subjects were the only ones emphasised in the lead up to the testing period.

An increase in whole class teaching, formal testing and emphasis on basic skills was reported. For example in science, there was an increased emphasis on memorising facts. The authors surmise that with too close a match between what is taught and what is tested, it is difficult to know whether students have conceptual understanding and the ability to apply those concepts or simply the ability to answer the test questions. High stakes tests may be affecting underlying learning.

Hancock and Kilpatrick (1993) report a substantial American study of mathematics teaching and assessment. They found that mathematics teachers stated that mandated testing had resulted in more whole class teaching, and fewer extended projects and interactive work. Teachers' test preparation practices gave rise to two concerns: the amount of time from regular instruction given to test preparation, and doubts about the educational value of that preparation. Research shows that repeated practice aimed strictly at the content of the test rather than the content domain of the test can increase scores without increasing student achievement. Classroom observation showed an average of 54 hours a year was spent preparing for external tests. Another study found that the percentage of teachers who spent more than 20 hours of class time on test preparation was three times higher for high-minority classes than for low minority classes. These reviewers found a negative impact of mandated testing on teachers who administer them, particularly in terms of workload and promotion associated with performance. There was evidence that standardised test scores played a major role in determining students' educational experiences. Test scores were the single most important factor in deciding on student placement. Young children experienced unnecessary frustration as they struggled with developmentally inappropriate standardised testing. Most teachers believed that standardised testing caused stress, frustration, fatigue, physical illness etc. A positive effect was stated to be that students became more serious about their classes. The authors conclude that the available research does not lead to the unqualified conclusion that mandated testing is having harmful effects on mathematics instruction. The picture is both more mixed and more indistinct. Since the research methods were questionnaires and interview questions there is need for further research which moves beyond self report to detailed classroom studies.

Jones and Whitford (1997) argue that the Kentucky Instructional Results Informational System (KIRIS), the accountability system introduced in Kentucky (which they describe as a bureaucratically controlled, high-stakes testing system), has distorted and undermined the original visions for effective curriculum, instruction and assessment practices. It turns the results of student performance assessments into a "school score" that the state then uses to determine sanctions or rewards for teachers and administrators. They state that it was too soon to see whether KIRIS had improved student learning.

In a change of geographical position but not of focus, Wilson (1996) comments on the Victorian government's standardised testing programme in Australia. She argues that the only legitimate purpose of assessment is to improve student learning and that it is questionable whether the Victorian tests do not do this because of their political agenda. Inevitably these tests send particular messages to the community about what is valued and what is significant in education.

Wilson claims that the Victoria tests are complex, and the process of recording and cross checking is difficult and time consuming, which has a negative effect on teaching and learning. The testing regime has increased teacher anxiety and stress, and ignores aspects of learning like

communication, investigation and cooperation. They claim that the assessment cannot be said to have improved student learning.

An aspect of the effect of curriculum on pedagogy explored in Plewis and Veltman (1996) and reported in Chapter 4, Section 4.2.2.2 was the time spent on different subjects. The authors conclude that the changed model of curriculum from 1988 did not have a major impact upon classroom time distribution. Science became a core subject in the new National Curriculum, but time spent on it increased only marginally, and well below the level recommended by the Dearing report on the National Curriculum and its Assessment. A mere two minutes a day were spent on computer work, and six minutes per day on geography and history, well below Dearing's recommendations. The authors could not comment on whether the changes had impacted on pupil achievement.

The introduction of a new national curriculum in the Netherlands in 1993/94 set out common objectives for 15 subjects for 12-15 year olds. An aim of the curriculum was that the knowledge acquired had to be true to life i.e. authentic and relevant to the students everyday lives. Students were also expected to play a more active role in the learning process. The curriculum required teachers to incorporate the characteristics of authentic pedagogy (Newmann and Wehlage, 1998) in their teaching. Roelofs and Terwel (1997) address the degree to which teachers use authentic pedagogy.

Three questionnaires were developed, covering four characteristics of authentic pedagogy:

- the construction of knowledge in complete task environments;
- the connection to the students' personal worlds;
- the value of learning beyond school;
- cooperation and communication.

In 1994 and 1996 a general questionnaire was administered to 89 teachers at 3 secondary schools. Eighteen teachers of English and mathematics were also interviewed on their use of textbooks in the context of authentic pedagogy. The results showed that none of the schools scored highly on the characteristics of authentic pedagogy. Teachers used examples taken from real life when explaining new subject matter to students, but students' own experiences did not figure as starting points in the lessons. The value of learning beyond school was addressed only as far as the textbooks permitted. Particularly striking was the difference in perception between teachers and students: the students were much less positive about the degree of displayed authentic pedagogy. The findings indicated that authentic pedagogy demands a big change in the teachers' role, including changes in the use of curricular materials and the development of new teaching strategies embedded in a supporting school organisation.

Summary of 4.3

This section has documented the complex effects on pedagogical approaches resulting from the interplay of curriculum and assessment. The discussion provides detail of the interactions represented in the analytical framework for the review (Figure 2). Some important findings from this section are:

- The common initial response to the development of a National Curriculum in England and Wales was positive, since a framework for learning had been provided. Teachers worked together to devise programmes for their classrooms.
- Performance type assessments can lead to enhanced teaching in particular curriculum areas, and provided better opportunity for many students to show best performance.
- The impact of the imposed assessment regimes in conjunction with the National Curriculum resulted in conflict, and the positive effects of the curriculum became undermined. Initially the national assessment employed SATs and these complex performance-orientated activities could be (with difficulty) incorporated into student learning because they had some coherence with teachers' previous practice. This positive

influence did not survive the many changes to national assessment and the high stakes consequences of ranking of schools and teachers. Pedagogical approaches focused on some aspects of the curriculum (with a marked trend towards more whole class teaching, less integration and more teacher talk), and teachers were concerned at the impact of this narrowing on student learning.

- The move to pen and paper tests and league tables led to streaming, subject based teaching, whole class teaching, greater class time preparing for tests, increase in transmission teaching and a move away from student centred learning.
- The increased subject areas in the curriculum in England and Wales required teachers to cut down student participation in order to get through the curriculum content.
- Innovations in other countries have also been strongly influenced by assessment processes.
- The teacher has had a central role in interpreting and implementing curriculum and assessment changes, as indicated in the analytical framework for the review (Figure 2).
- The amount of instructional time given to curriculum areas was found to be directly related to the national testing, therefore the prescribed assessment programme contributed to a narrowing of the curriculum.
- Teachers of high minority classes used more time (up to three times as much) on test preparation than for low minority classes, thus reducing the curriculum even further for these students.
- Most of the research data on assessment in the USA and the UK appears to be on teacher report data. More classroom research is needed on the interaction between curriculum, assessment and student achievement to develop a greater understanding of this complex area.

4.4 The effects of curriculum on learning and social outcomes

The discussion so far in this chapter has indicated a complex interaction between curriculum, assessment and pedagogy, which through the intervening role of the teacher comes to affect students' learning and social outcomes (see Figure 2). It is difficult, if not impossible, to find research evidence for direct influences of curriculum on student learning and social outcomes.

In the past, research linking curriculum and learning has been largely limited to studies published within journals dedicated to particular curriculum subjects. These researchers tend to investigate learning as it applies to a particular subject. Working along these lines, Stodolsky, Salk and Glaessner (1991) investigated the way students thought about learning in relation to maths and social studies by interviewing 60 fifth grade students. The students' responses showed that they saw considerable differences between subjects. For example, a positive conception of maths was linked to the students' success or ability to do the work, while a positive conception of social studies was aligned to student evaluation of the interest level of a lesson or activity.

Rather than looking at the differences between subjects, some researchers have analysed classroom activity by looking at aspects of the academic task. By analysing classroom activities in terms of the generic task researchers hoped to gain a better understanding as to ways teacher could structure tasks to enhance student learning. Reviewing 196 empirical and theoretical studies of teaching and learning, Doyle (1983) suggests that there are three aspects of the academic task: the product (eg. an essay), the operation used to generate the product (eg. critical thinking, memorising) and the resources available to students (eg. a model essay). Students acquire knowledge and practice operations in the process of completing these tasks. Doyle (1983) concludes his meta-analysis, asserting that the way teachers introduce and manage tasks, together with the expectations they communicate to students has a profound effect on what the students gain from completing the task.

Using the framework developed by Doyle (1983), Doyle and Carter (1984) conducted an intensive case analysis of the academic tasks completed in three English classes taught by the same teacher. Doyle and Carter (1984) found that although the teachers attempted to engage the students in tasks

that demanded high level cognitive processing, the students' response was to consistently demand more and more direction and specificity in order to reduce the ambiguity and risk involved in completing such a task. The result of this interaction was that although many tasks initially demanded a high level of cognitive processing they gradually became downgraded to low-level tasks. Doyle and Carter (1984) argue that this result raises crucial question about the achievement of high level curriculum objectives in the classroom, and what can be done to support teachers to translate high level curriculum objectives into tasks that are manageable from the point of view of students and teachers.

Brophy and Alleman (1992) developed this notion further in an analysis of 43 mainly theoretical studies relevant to the themes of curriculum, classroom activity and student learning, coupled with an analysis of curriculum materials. Brophy and Alleman(1992) suggest that a curriculum should not separate skills and knowledge. Instead it should be aimed at engaging students in higher order thinking so that they can develop understandings that can be applied to life outside the classroom. Informed by their analysis Brophy and Alleman(1992) have developed a list of principles that propose a set of criteria to be used in designing what they describe as an "ideal curricula":

- curriculum development should be driven by significant long-term goals, not content coverage lists;
- content should be organised into networks around important ideas, and any teaching of these ideas should be aimed at understanding and application to life outside the classroom;
- activities are not ends in themselves but are means for assisting students achieve major curricular objectives;
- knowledge and skills should be integrated and should be consistent with long-term curricular objectives;
- strands or sets of activities will serve different purposes within a curriculum, and these functions will evolve as strands develop;
- activities should be assessed in terms of costs and benefits;
- the activities included in the proposed curriculum should be those that cognitively engage learners so that they restructure ideas and undergo conceptual change as they complete activities;
- the success of any activity in moving students to thoughtful engagement will depend not only on the structure of the activity, but the way the teacher implements the activity, and engages the students in dialogue before, during and after the period in which the students complete the activity.

Brophy and Alleman (1992) conclude that although activities are the basic building block of classroom action, they are also the means through which larger curricular goals are achieved.

Oakes, Wells, Jones and Datnow (1997) conducted a three year longitudinal case study of ten secondary schools that were involved in detracking students so that classes contained a heterogeneous mix of students in terms of ability. The researchers interviewed students, teachers, parents, administrators and community leaders as well as conducting observations in classrooms and various school meetings. The results of this study showed that when students were in classes of perceived similar ability, white middle class students tended to dominate the higher stream classes whereas minority and low-income students were clustered in the lower streams. Schools that started detracking reform experienced high levels of resistance from parents of students in the upper streams and most teachers, who tended to see student ability and intelligence as innate and heritable traits. However, once detracking had taken place, many teachers found that they had to rethink their conceptions of ability and intelligence as they noted that alterations in the way they delivered the curriculum had a profound effect on student achievement. The researchers conclude that detracking represents more than a rearrangement of students grouping; it is a political endeavour that challenges entrenched cultural beliefs and ideologies. Furthermore, it raises questions about traditional delivery of resources whereby those students in the upper streams are advantaged and students in lower streams disadvantaged.

Kozol (1991) spent two years visiting a variety of schools situated in inner city and less affluent suburbs across the America, observing classes and talking with teachers and students. Although he provides little detail of the scope or breadth of these observations and interviews, the main thrust of the resulting narrative account is to describe the wide gulf between the quality of education offered at these schools compared to that provided in schools situated in more affluent areas. Teachers in the 'poor' schools are generally unmotivated, resources are scant, and the curriculum is barely covered. Kozol (1991) posits that it is the way minority and low-income students are treated in the school system rather than their culture or socio-economic status that has the greatest effect on their academic achievement .

Further discussion of this topic will be deferred to Chapter 7 after consideration of assessment in Chapter 5 and pedagogy in Chapter 6.

General summary

Overall this chapter has indicated that:

- There are no firm research-based conclusions that can be given as to whether a curriculum should be mandated or open at the national level or to what level of specificity learning outcomes should be defined.
- There is no clear evidence that the existence or absence of a national curricula leads to improved student performance.
- All countries in the OECD including those high-performing countries in TIMMS are revising their curricula to include more practical work, relevance and motivation.
- The area of progression is one of the least understood areas, yet this is crucial for designing curriculum and enhancing student achievement. This area requires much more research in all curriculum areas.
- There is a dearth research evidence to understand the complexity of curriculum, assessment and pedagogy and its influence on student achievement at the classroom level.
- Evidence is available that shows that when critical thinking skills and content are included in a school curriculum student performance can be raised.
- High stakes assessment can limit the classroom curriculum for students, particularly for lower achievers and minority students.
- No matter what is included in a curriculum at the national level, it is the classroom curriculum that impacts on student achievement, and there are very few classroom studies that provide insights into this.
- More classroom-based research needs to be undertaken on the interaction between the different levels of curriculum, types of assessments, pedagogical approaches and student achievement.
- The types of assessment regimes that are put in place influence curriculum time on subjects.
- If curriculum reforms and assessment reforms take place at the same time, how much change was due to which variable is unknown and will need further exploration.

Chapter 5: Assessment

Promoting children's learning is a principal aim of schools. Assessment lies at the heart of this process. It can provide a framework in which educational objectives may be set, and pupils' progress charted and expressed. It can yield a basis for planning the next educational steps in response to children's needs. By facilitating dialogue between teachers, it can enhance professional skills and help the school as a whole to strengthen learning across the curriculum and throughout its age range (DES, 1987).

It is deceptive to speak of 'assessment' in the singular, as if it were a uniform activity perfectly controllable by a single agent. Assessment intervenes in the educational system in many ways, especially summative assessment. It is often a final verdict: it can determine categorical decisions about students or sanctions upon them; and it often shapes the judgments of parents, teachers and policy-makers. It links the world of learning with the society outside. The positive or negative value that it attaches to a young person may be hard ever to alter. There is no such thing as simple assessment of a student's practice of a discipline, or of a student's acquisition of knowledge in a discipline, or even of a student's awareness of acquiring such knowledge. Assessment is multiple. It has to embrace the whole work of learning and it has to relate to the whole work of teaching, and then to assess the effectiveness of this combined work (Black and Atkin, 1996).

Introduction

Assessment of students can be for several purposes, namely:

- to give feedback to students about their learning and progress.
- to give feedback to teachers about the effectiveness of their teaching and curriculum planning.
- to report on student learning and progress to parents, caregivers and the school management.
- for the award of national qualifications and accreditation of individuals and institutions.
- to evaluate the education system, through public accountability of institutions and teachers.

Assessment for all the above purposes has the potential to intervene (either positively or negatively) in learning, pedagogy and the curriculum. These interventions will be discussed in this chapter. Assessment practices in schools have the ability to attach to a young person a positive or negative judgment and value which may be difficult to alter (Black and Atkin, 1996).

This chapter will first define key terms in assessment (section 5.1), list some main issues in assessment (section 5.2), review the literature on assessment (5.3), the effects of assessment on pedagogical approaches (5.4), and the effects of assessment on learning and social outcomes (5.5).

5.1 Key terms in assessment

The literature distinguishes two key types of assessment: formative and summative assessment.

The New Zealand Ministry of Education (1993a) defined **formative assessment** as:

A range of formal and informal assessment procedures (for example, the monitoring of children's writing development, anecdotal records, and observations) undertaken by teachers in the classroom as an integral part of the normal teaching and learning process in order to modify and enhance learning and understanding. (Ministry of Education, 1993a)

Other statements about formative assessment are:

Formative assessment has been defined as the process of appraising, judging or evaluating students' work or performance and using this to shape and improve students' competence. (Gipps, 1994)

The distinguishing characteristic of formative assessment is that the assessment information is used, by the teacher and pupils, to modify their work in order to make it more effective. (Black, 1995)

Formative assessment is feedback to the teacher and to the student about present understanding and skill development in order to determine the way forward. (Harlen and James, 1996)

Formative assessment is defined as the process used by teachers and students to recognise and respond to students' learning in order to enhance that learning, during the learning. (Cowie & Bell, 1996)

Within this review, diagnostic assessment, which aims to diagnose student weaknesses, is viewed as a part of formative assessment. These definitions emphasise that assessment is formative in its function only when action is taken which is intended to improve student learning. They also indicate the need for the involvement of both teachers and students in the process of formative assessment. Assessment to support learning is also referred to as educational or educative assessment. Gipps (1994) identified Glaser (1963) as the first to propose attention be given to 'educational assessment'. She argued that a shift from assessment to prove learning to assessment to improve learning was signalled by Glaser's suggestion of a move from norm to criterion referenced assessment.

Summative assessment is used for the purposes of describing learning achieved at different times for the purposes of reporting to parents, other teachers, the students themselves, and in a summary form to other interested parties such as school governors, school boards or accreditors of national qualifications (Harlen and James, 1996). Summative assessments may also be used for public auditing and accountability of institutions and individuals, and for government policy review. Hence, summative assessment in New Zealand includes assessment for national testing, such as in the National Education Monitoring Project; assessment for national qualifications in the post-compulsory sector; and assessment for school reports to parents and care-givers.

Summative assessment may take place on one occasion (such as by external examination at the end of an extended period of time), or be continuous (for example internal assessment by the teacher throughout the school year). Continuous summative assessment is a weak form of formative assessment.

The term 'assessment regime' is used here to refer to the total package of assessment policies and practices at the state or national levels. In New Zealand, this would refer to the mandated assessment requirements of assessment of curriculum outcomes for reporting to parents and care-givers; assessment for national qualifications (for example, Achievement 2001); and assessment for public accountability (the information required by the Education review office, NEMP and the proposed national testing of primary schools).

5.2 Main issues in assessment

This section provides a preliminary commentary on research into assessment and its affect on learning, pedagogy and curriculum.

Black (1995) in a commentary on an extensive review of assessment describes the functions of assessment as:

- direct assistance to learning, wherein assessment information is used, by both teacher and pupils, to modify their work in order to make it more effective.
- the certification of individual students, and
- public accountability of institutions and teachers.

The first function is broadly formative assessment and seen as a major opportunity to enhance classroom learning. The second and third functions are broadly summative.

Black notes that these functions are interactive; whilst the accountability function does not require data on individuals, in practice it is often operated using certification data from the assessment of all students; most of the investment in assessment and testing, whether in practical operations or in research and development, has been devoted to the certification and accountability functions, to the neglect of the formative.

The result is that external testing can dominate classroom work and so distort teaching that the conditions for good formative assessment do not exist. Their overview is that narrow external testing can have detrimental effects on teaching because: learning then follows testing in focusing on aspects that are easy to test, lowering the cognitive level of classroom work; pupils often have to work at too great a pace for effective learning when external testing dominates; creative, innovative methods and topical content are harder to include in classroom practice. These trends are argued to demotivate many pupils, whilst rewarding pupils who work in the narrowly constrained ways that test success requires.

Black argues that when summative practice dominates, teaching and learning are influenced in the following ways:

- in summative testing the need for a single overall result means that quite disparate data (e.g. for practical and for theory) are combined in an arbitrary way: formative assessment need not do this since it is concerned with the patterns of learning needs of individual pupils;
- summative work is seen to require standards of uniformity and reliability which hamper attention to the individual needed in formative work, where different pupils can be treated differently;
- summative purposes require documented evidence for results and so adds to teachers' workload, whereas formative work calls for action on the data rather than its storage.

Further issues arise from concerns:

- about the use of data collected for formative purposes being used for summative purposes, and vice versa.
- at the lack of robust evidence for the advantages and disadvantages of assessment tasks such as portfolios, observations of activities, and multiple choice tests.
- at perceived mismatches between the achievement objectives in curriculum documents and the assessment objectives in national documents.
- at the possible introduction of national testing into primary schools, in terms of the potential to distort classroom curriculum and pedagogy.
- about the development of assessment practices which would be equitable for students of different race, social class and gender.

The remainder of this chapter provides a commentary on the research identified in this review which seeks to provide information on these issues.

5.3 Research on assessment

Section 5.3.1 discusses the issues of validity and reliability of assessment processes raised above. Section 5.3.2 outlines research which describes significant gains in learning achievement from formative assessment procedures, and records the substantial changes to common classroom practice implied in a change to formative assessment in support of teaching and learning. The issue of teacher professional development is then raised.

5.3.1 Validity and Reliability

The issue of the validity and reliability of assessment procedures has been discussed in the literature reviewed. Black (1995) argues that external written tests command more public

confidence than they should, since the limits to their validity and reliability are not understood. He cites Messick (1989) on validity:

Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the **adequacy** and **appropriateness** of **inferences** and **actions** based on test scores or other modes of assessment.'

He notes that the scope of these inferences and actions provides a different perspective on the validity of external tests. For example, the uses to be made in teaching and the backwash effects on teaching need appraisal. On this basis if an assessment could be judged to reinforce styles of teaching which are inappropriate for the aims of education, its validity could be prejudiced on these grounds alone. Black's analysis is that the closer an assessment activity can come to the actual activity to which its results are to be considered relevant, the more it is likely to satisfy validity criteria. Black expresses concern that the limited reliability of formal tests is not appreciated, and in many national systems is not measured. This is a concern since for a variety of reasons pupils often underperform in formal tests. He cites in evidence Gauld (1980) who found that pupils in science tests often misread the demand of a question; seemed incompetent because of a single slip in a complex process; failed to use what they knew because they judged it irrelevant; and could be marked down when the marker failed to understand the thinking behind non-standard responses.

Murphy (1996) has a complementary comment on the IEA testing process. The science tests are shown to examine both concepts and variables, so that failure might result from conceptual or procedural difficulties, or both. The mathematical demand of some science items has also been found to depress performance. In Murphy's view the TIMMS material has value for uncovering difficulties in science and mathematics, rather than for developing national policy.

Bracey (1996) argues that the TIMSS reports cannot be compared with reliability because too many variables are ignored. He cites the case of Japan where he claims that unquestioning obedience and extreme discipline in schools, and the success of the juku after school coaching classes which most students attend, skews the results.

In a further substantial commentary on the accuracy and reliability of the IEA tests, Kellaghan (1996) is concerned at their use for developing public policy since:

- different curricula are associated with different patterns of achievement. Therefore the interpretation of achievement difference between countries is problematic.
- sampling procedures are not equivalent.
- the measures are mainly of recall and do not include aspects such as analysis, decision making and critical thinking.

Kellaghan argues that while educational achievement may intuitively seem to be an important factor in determining a country's economic activity, the precise relationship between educational achievement and economic activity or growth rate merits further investigation.

Madaus and Kellaghan (1992) reported earlier concerns about the use of high stakes assessment. Some schools, they claimed, deliberately manipulated various aspects of schooling to enhance the school's test results, such as dropping out those students unlikely to perform at a 'passing' level. Madaus and Kellaghan claim that the evidence from their review of assessment research is that as long as schools rely on outcome measures, which are usually standardised test scores, then there will be continued difficulty in judging the value of curricula.

Jones (1996) reports that Standard Assessment Tasks (SATs) administered in Britain had major weaknesses, and that many (particularly in language assessment) provided little information of any value. The SATs were claimed to rely too much on teacher trust, and to lack essential validity and reliability. Jones questioned the administration costs of SATs (NZ\$90 million), and argued that it was unwise to rank schools on the basis of performance on SATs.

Lewis (1996) is also concerned at the use of standards tests and tasks with primary school children, particularly those with special needs, since the process of combining data masked attainment in other areas.

In a detailed empirical study of results in schools in England and Wales at Key Stage 1, Strand (1997) reports an attempt to measure the progress of students in 1995 in terms of a value-added analysis. Evaluation by schools in this manner is argued to be important in identifying effective practice and in initiating change. The author concludes that while this valued-added analysis identifies differences between schools in the progress made by their pupils, there is some way to go to explain the difference. Schools which obtained high scores were not necessarily the most effective in this measure of promoting pupils progress.

Becker (1998) discusses standards and testing in the US and argues that the enormous cost of developing and administering standardised tests produces multiple choice tests which are extremely limited in the amount of information they can provide about students, and that moving beyond multiple choice or fixed response tests is very expensive. According to Becker there is wide spread evidence of cheating by teachers, students and administrators on state and national standardised tests in the United States, pointing out that in Kentucky 60 schools have been investigated since 1993. Becker also claims that standardised testing is likely to lower standards by lowering assessment to the lowest common denominator. Moreover, in a pluralistic nation like the United States, it is difficult to reconcile factions who claim tests are unfair to particular minority groups. Shifting the focus to outcomes, Becker cites research by Card and Krueger (1996) which shows little relationship between standardised test scores and productivity.

Poston (1997) reviews current assessment finding in curriculum audits and claims that assessment in schools tends to be broad and large in scale, customarily using paper and pencil tests, that data gathered is seldom disaggregated, and inferences are not drawn from data. The author argues that assessment needs to have clear goals, clear targets, rely on proper methods and samples, and include measures to control bias. Currently the scope of assessment is often inadequate and the types of assessment used in schools are narrow.

In summary, the research reviewed indicates that external testing can lead to a narrowing of classroom experiences, manipulations of student experience to achieve higher scores, and higher compliance costs without corresponding gains in student achievement on curricula outcomes. The research raises particular concerns at the quality and usefulness of the findings of large scale assessments such as IEA and TIMMS in making policy on assessment.

The research reviewed also documents problems about high stakes national assessment, due to its effect on pedagogy, on some student groupings, and its apparent encouragement of minimal competency. Nevertheless, it should be remembered that some advocates of high stakes national assessment actually support teachers teaching to the test, for that in itself focusses teachers on the content required for the test. In this sense, high stakes testing is not, by definition, negative.

The above commentary on validity and reliability focuses on summative assessment. For formative assessments, the criteria of validity and reliability need also to be reformulated. Black (1995) argues that the validity of formative assessment is to be appraised in relation to its effectiveness in improving learning. Cowie and Bell (1996) discussed validity of formative assessment in terms of trustworthiness, appropriateness and equity.

5.3.2 Assessment to support teaching and learning

In much of the research and literature on assessment Crooks' (1988) meta-analysis of more than 240 studies relating to evaluation practices is mentioned. Crooks (1988) analysed these studies by looking at the variety of evaluation practices used by teachers and their effect on students learning.

The analysis revealed that although classroom evaluation has a profound effect on student learning, it is an aspect of teaching practice that receives little attention in teacher education courses or professional development programmes. The classroom evaluation practices that teachers tend to use place emphasis on surface learning and rote memorisation, rather than deep learning or high level cognitive processing even though research shows that students quickly forget isolated details and are more likely to remember and be able to apply information if it is embedded in a wider framework of learning. Crooks (1988) also found that there was an emphasis on the grading function of evaluation, and that its usefulness as a way of assisting students to learn was neglected. The most effective evaluative practices were those that were: specific to each student and learning situation; were regularly timed so that the students could use the feedback to alter their approach; were structured in such a way that they suited the assessment goals. Teachers who set high but attainable standards gained the best results from students (Crooks, 1988).

In an attempt to evaluate the knowledge teachers' possess about educational assessment, Plake and Impard (1997) conducted a study in which 555 teachers from elementary, middle and high schools across the U.S.A. completed a 35 item multiple-choice test. The results of the national survey showed that, on average, the teachers scored 66%, which the researchers suggest, indicates the overall inadequacy of teacher assessment literacy. Given that there is a move for teachers to use more authentic assessment strategies in the classroom and that these strategies require even more knowledge of assessment, Plake and Impard (1997) conclude that much more attention needs to be placed on developing teacher knowledge in this area.

Shepard, Flexen, Hiebert, Marion, Mayfield and Weston (1996) worked with 13 third grade teachers from three schools over one year in an attempt to encourage the teachers to use performance assessments (open-ended summative assessments) as a regular part of their instructional programme. The design of the project was based on the assumption that the use of performance assessments as opposed to multiple-choice summative assessments provides a more accurate measurement of student learning and improves teaching and learning. The researchers worked with the teachers, encouraging them to meet curriculum objectives by using open-ended, problem oriented tasks. At the end of the year, the achievement of students was measured using a standardised test, and was compared to the achievement of students at the three control schools. That there were no perceptible differences between the achievement of the students at the experimental and control schools suggested that implementing performance assessments had not improved the teaching or learning. Suggesting reasons for this result, the researchers noted that the teachers in the experimental group had not adapted their classroom instructional patterns as the researchers had hoped, indicating that the level of professional development included in the project had not been great enough. Shepard et al. (1996) concluded that while authentic assessment techniques might be an element in creating more effective learning environments, they are not an answer in themselves and that the role of teachers in implementing change should not be underestimated.

Section 5.1 raised the issue of the value of formative assessment in the teaching and learning environment. A special issue of the journal *Assessment in Education* (March 1998) devoted to *Assessment and Classroom Learning* explored this issue. A review paper by Black and Wiliam was followed by invited commentaries from other researchers in the field. The paper and the commentaries contain a major overview of the relevant research. Black and Wiliam's paper describes a search from 1988 until the time of publication in which 681 publications were scanned and about 250 became the basis of their paper. The other authors in the journal refer to many more publications in their separate commentaries. In summarising their review Black and Wiliam (1998) make two comments very relevant to this review. The first is about the process of their review. Their analysis of the factors which combine to determine the effects of any classroom regime reports that the systems are complex and multivariate. They note that most of the studies they reviewed did not attend to some of the important aspects of the situations being researched such as:

- the assumptions about learning underlying the curriculum and pedagogy;
- the rationale underlying the composition and presentation of the learning work;

- the precise nature of the assessment evidence revealed by the learner's responses, the interpretative framework used by the teachers and learners in responding to this evidence, and the learning work derived from these interpretations;
- the perceptions and beliefs of learners and teachers about the assessment processes and intentions behind them;
- issues related to race, class and gender, which appear to have received little attention;
- the extent of artificiality of the research study and the effects of this on generalisability.

Despite these concerns, which relate to the complexity of the research setting rather than to inherent deficiencies in the studies, they conclude that

For public policy towards schools, the case to be made here is first that significant learning gains lie within our grasp. The research reported here shows conclusively that formative assessment does improve learning. The gains in achievement appear to be quite considerable, and as noted earlier, amongst the largest ever reported for educational interventions. As an indication of just how big these gains are, an effect the size of 0.7, if it can be achieved on a nationwide scale, would be equivalent to raising the mathematics attainment score of an 'average' country like England, New Zealand or the United States into the 'top five' after the Pacific rim countries of Singapore, Korea, Japan and Hong Kong. If this first point is accepted, then the second move is for teachers in schools to be provoked and supported in trying to establish new practices in formative assessment, there being extensive evidence to show that the present levels of practice in this aspect of teaching are low, and that the level of resources devoted to its support, at least in the UK since 1998, has been almost negligible. (Black and Wiliam, 1998, p 61)

They further comment on this page that 'there is enough evidence in place for giving helpful guidance to practical action' and 'despite the existence of some marginal and even negative results, the range of conditions and contexts under which studies have shown gains can be achieved must indicate that the principles that underlie achievement of substantial improvements in learning are robust.' Black and Wiliam (1998) also note that 'the changes in classroom practice that are needed are central rather than marginal, and have to be incorporated by each teacher into his or her practice in his or her own way. That is to say, reform in this dimension will inevitably take a long time, and need continuing support from both practitioners and researchers' (p. 62).

That this reform is important for minority students is strongly indicated in the study of the impact of mandated standardised testing on minority students in mathematics and science (Lomax, West, Harmon, Viator & Madaus, 1995) which examines the six most widely used test batteries in the US. The test batteries were evaluated to determine their emphasis on thinking, conceptual knowledge, and procedural knowledge in mathematics and science. Survey data were obtained from 2,229 mathematics and science teachers of high percentage and low percentage minority classes, and interviews were conducted with 289 urban educators. The tests were found to fail to adequately sample higher order thinking, high level conceptual or high-level procedural knowledge in either subject. Teachers of high percentage minority classes were found more likely to report negative impacts of standardised testing on teacher practice and minority student achievement in mathematics and science. The study concluded that having more instruction on preparing for mandated testing adversely affected the quality of instruction in high percentage minority classrooms. Hence, increasing the amount of, and the stakes associated, with mandated testing puts minority students in these classrooms at a greater risk of failure.

Black (1995) provides a focused commentary on the value of formative assessment as a support to teaching and learning and the changes which would be implied in the classroom. He argues that formative assessment is intimately connected with the process of teaching and learning, requiring integration of testing and instruction, and a clear conception of the curriculum, the goals, and the process of instruction. In terms of the curriculum he notes the problems associated with presenting

an idea which is important in many contexts in abstract isolation. Learners also have difficulties when complex skills are taught and tested in a fragmented manner.

Black (1995) cites the research of Simon et al (1994) about the need for feedback given by formative assessment to reflect important aims; be criterion-referenced; and to reflect well-founded knowledge of the sequences in which pupils learn. This requires management of a curriculum plan constructed by teachers for themselves, and enhanced when teachers share assessment information and practices. Teacher knowledge is then crucial.

The role of the pupil is seen as pivotal in formative assessment since many pupils may have acquired habits of doing just enough to get by, or have ceased to believe that they can be competent at the subject. To overcome these difficulties pupils need to understand the frame of reference of the teacher, and the model of learning which gives meaning to the criteria that are reflected in their assessment. This understanding can build up the pupil's confidence in their understanding, and their confidence that they can do something about their difficulties.

A contribution to pupil encouragement would be the translation of curriculum aims, into language that all pupils can understand, and down to a level of detail that helps them related directly to their learning efforts. Targets would need to be attainable in the short term, and adequately modest in relation to the learners' prospects of success.

The capacity of pupils to judge their own work is of fundamental importance in learning. Zessoules & Gardner (1991 as cited in Black, 1997), describing the Harvard Project Zero, emphasised that self-assessment at the point of learning is a crucial component for developing complex understanding through reflective habits of mind, so their project tried to set up conditions to ensure that students were active and thoughtful in their own assessment. Others have argued that metacognition, by which they mean awareness and self-direction about nature of their learning work, is essential to pupils' development in concept learning (see Brown, 1987; White and Gunstone, 1989, cited in Black, 1997).

Clearly, pupils' involvement can make it more feasible for teachers to carry through a programme of formative assessment. However, this involvement also changes both the role of the pupil as learner and the nature of the relationship between teacher and pupil, making the latter shoulder more of the responsibility learning. Thus improved formative assessment can lead to changes which are of much wider significance - changes which should be a powerful help with pupils' personal development and which should also be part of any programme to help them to be more effective learners.

The changes to classroom practice implied in Black's analysis could not simply be added to existing schemes of work. He argues that in their assessment responsibilities, teachers have to reconcile the learning needs of their pupils, which it must be their first concern to serve, with pressures to obtain good results in statutory national tests and expectations that they will work within a framework of school policies and respond to parental expectations. These requirements are often inconsistent. If teachers are to make radical changes to their classrooms they will need to support one another, and receive sympathetic understanding from those outside schools. This process requires teacher professional development processes which match the many demands of the innovation outlined above.

Summary

This section provides a major contribution to this review which fits with the analytical framework (Figure 2). Black and Wiliam (1998) and the associated literature present a convincing argument for formative assessment as a means to improve teaching and learning. Black's commentary on formative assessment reflects on the many interactions between curriculum, assessment and pedagogy, and highlights the central role of the teacher and the learner. Any change towards classroom implementation of formative assessment is seen to be complex, involving interaction between curriculum, assessment and pedagogy, as is outlined in the review framework. This

section also indicates the need for teacher professional development to support changes to the system.

5.4 Effects of assessment on pedagogical approaches

This section reports the research related to the effects of different assessment approaches on pedagogy, noting that the interaction is strong. The concept of high stakes assessment provides a focus for the discussion. When national (and other systematic) curricula and associated assessment regimes have been devised there can be a substantial effect of one on the other, often contrary to the spirit of the curriculum or the assessment. This review found supporting evidence for problems being created in the classroom by external monitoring. Most comment came from the UK because the introduction of national testing accompanied the introduction of a National Curriculum, and where a considerable body of research has focused on the consequences for pedagogy.

The tone of the research commentary comes in a major review of key articles on assessment by James and Gipps (1998). They argue for a reconsideration of the types of assessment promoted by national policy; that education for the 21st century should place emphasis on higher order skills and 'deep learning' while not neglecting the basic skills. Their main arguments (with acknowledgment to the cited reviews of Broadfoot, 1998; Crooks, 1988 and Black and Wiliam, 1998) are:

- assessment provides learning motivation by giving a sense of success in the subject (or demotivation through failure) and a sense of self-confidence as a learner;
- assessment through feedback on success helps students (and teachers) decide what it is important to learn;
- assessment helps students learn how to learn by encouraging appropriate learning styles and strategies, by inculcating self-monitoring skills, and by developing the ability to retain and apply knowledge, skills and understanding in different contexts;
- assessment helps students learn to judge the effectiveness of their learning by evaluating, consolidating or transforming existing learning, and by reinforcing new learning.

The authors acknowledge that the range of learning from deep to surface requires assessment strategies which will match this range, for example multiple choice tests for recall of facts (surface) and essays to explore knowledge of the structure of the discipline and concepts and principles, applying knowledge and linking it to real life situations (deep). They accept that use of a particular assessment approach will not necessarily move teaching and learning in the hoped direction, but also point out that remaining with narrow approaches will not either.

Their commentary on mandatory national testing is that it has come to focus mainly on those types of learning which require the memorisation of facts and the display of basic skills, thus privileging one kind of learning and damaging the higher skills promoted in the curriculum. Their recommendations are that:

- national monitoring should be based on the results of both short tests and more extended tasks, to represent the full range of learning objectives;
- national tests which assess retention and recall of factual information should be balanced by more extended assessment tasks which provide opportunity for students to demonstrate higher order skills in contexts other than those in which they have learnt them;
- a proportion of the resources used to develop national tests should be used to support the professional development of teachers as assessors for the kinds of learning appropriate to the curriculum and future needs;
- reporting to parents and other stakeholders should be clear about the students' different achievements and difficulties in relation to the different kinds of learning (thus having difficulty with spelling does not imply poor achievement at English).

As already noted in Chapter 4.3 the national monitoring process is acknowledged to have stimulated some positive effects on pedagogy. There is evidence from the National Assessment in

Primary Schools evaluation project (Gipps, 1996b) that for teachers in classrooms for children at Key stage 1 (7 years) and at age 11 year the national assessment resulted in:

- increased levels of discussion and collegiality in the early stages of the implementation when teachers were negotiating meanings for the SATs (standard assessment tasks) and devising moderation procedures.
- raised professionalism of the teachers involved in leading assessment training and policy development. The teachers became more knowledgeable on assessment.
- greater care in planning, closer observation of children and a more detailed understanding of individual progress in the opinion of over half of the 8 principals in the study.

A further commentary on formative assessment comes from Sadler (1998). Key points made in this paper are:

- Formative assessment is effective in all educational settings.
- Grades do not deliver as much formative effectiveness as tailored comments.
- Quality is crucial.
- Students need to be trained in how to interpret feedback.
- Any movement towards feedback-enhanced learning conditions must be carried out long enough for the new procedures to be viewed by the learners as normal and natural.
- Teachers can only be effective at formative assessment if they know both sides of the operation - how students learn and the subject area.

Sadler notes the need for further research which includes further analytical work, and comprehensive empirical research which is ecologically valid. The difficulty for this research will be that results may be masked by other factors.

Other issues explored are more complex. A number of research studies conclude that teachers are using restricted pedagogy as a consequence of 'teaching to the test.'

Silvernail (1996) reports interviews with 50 teachers, head teachers, educational officials and policy researchers in 1993. The majority of teachers had fundamentally changed the way they teach, what they teach and how they assess their students as a result of the implementation of the national curriculum and assessment processes in England and Wales. Although teachers made considerable efforts to implement the intended curriculum which they saw as broad and balanced, it appeared that a piecemeal approach was narrowing these broad aims. The amount of instructional time became directly related to the anticipated national testing which diluted the classroom curriculum by limiting the time available to explore new areas. The SATs (pen and paper version) were not integrated into teaching and the introduction of national league tables resulted in temporary and permanent exclusions increasing by 66%.

The research cites three other studies which have documented how high stakes traditional tests result in narrowing the curriculum, and encourage teaching to the test.

Gipps (1996b) provides a somewhat different perspective in comments that the teachers were not so much teaching to the test as teaching areas of the curriculum so that their children could carry out assessments on those topics. Since there is an imposed curriculum as well as the imposed assessment, the teachers are consequently teaching to the curriculum as well as to the tests. Furthermore the rolling programme of attainment targets precluded teachers from too narrow a focus in preparing pupils for tests. All the above came about, not because of the national testing but because of the demands of teacher assessment and performance type activities in the SATs which were in the curriculum. Gipps comments that to raise real standards of teaching and formative assessment (which in turn supports teaching and learning), more than imposed external traditional tests is required.

Madaus (1991) asserts that educational policy makers should more carefully consider the effects of implementing 'high stakes' national testing before students are subjected to such a testing regime.

External assessments become 'high stake' when the results are published and passing the test is considered an important achievement. The results of high stakes external assessments are often used as an informal yet powerful measurement of teacher and school effectiveness. Madaus (1991) presents a catalogue of advantages and disadvantages of implementing a national testing regime. Included in the advantages are that national tests or examinations are a relatively objective and impartial means of distributing educational benefits because of the equality of access of the test to all students, and that they provide a degree of homeogeneity in national educational standards. On the other hand, Madaus points out that national tests often have the effect of limiting curricular variety because teachers "teach to the test", and that the results of tests are often used as a means of ranking students which has the effect of disadvantaging minority and low-income students.

The results of an extensive research study by Smith and Rottenberg (1991) provide some support for the cautions ventured by Madaus (1991). In this study, researchers spent 15 months in 2 elementary schools observing testing and related activities in classrooms, and interviewing teachers to see the effect that external 'high stakes' test had on decisions about what to teach, how to teach it and how to organise schools. The researchers found that external tests significantly reduced the amount of time for teaching subjects outside the realm of the tests. The teachers focused on teaching activities that were linked to the test or were similar in structure to those included in the tests, and neglected to teach activities or subjects that were excluded from the external tests. Although the teachers felt that test scores did not accurately measure student ability or achievement, scores were used to place students into ability classes. Overall, the teachers thought that external tests negatively affected young students or students who struggled to accomplish skills needed in test situations and that external tests had a negative impact on their own performance as teachers (Smith and Rottenberg, 1991).

Shepard (1991) also questions the value of implementing 'high stakes' national testing, noting that they have often been implemented as a knee-jerk reaction to a perceived dip in student achievement. A review of 25 relevant studies points to the fact that these regimes have caused teachers to focus more on teaching basic skills and rote memorisation of facts and that this emphasis is at the expense of higher order thinking. Shepard (1991) maintains that for summative assessment to achieve the result of improving students' learning opportunities, they must not be of the multiple-choice variety, should centre on higher order thinking and their implementation should be linked to teacher professional development.

Dorn (1998) reviews the effects of instituting national testing programmes, based on the results of close to 130 relevant research articles. He asserts that although research clearly indicates that national testing distorts teaching and does not provide a clear picture of student achievement, it nevertheless remains popular amongst legislators and the public because it provides some degree of certainty in an area that is characterised by uncertainty. In the long term Dorn (1998) argues that high stakes standardised testing will result in a narrowing in both teaching focus and the way schools are judged as institutions. Moreover, because schools and educators are judged by national tests, there is an incentive for teachers to try to exclude students with disabilities from their statistics. Indeed, teachers in Tennessee (a state in the U.S. that has implemented high stakes testing) have made just this plea. This phenomenon indicates that high stakes testing regimes perpetuate the pattern in which minority or low-income students receive a diminished share of educational resources and opportunities (Dorn, 1998).

Brown et al (1996) provide further research on the impact of national assessment on teaching. They note that the national testing as used in the UK is high stakes (Madaus, 1988). The research reported is from the National Assessment in Primary Schools (NAPS) project. This focused on key stage 2 testing and 31 schools in 4 very different LEAs between 1990 -96. Extra data was collected in 5 case studies. The article focuses on the impact of key stage testing on primary schools and classrooms, on the use of the information by secondary schools and on the interpretation of these results when they were made public. The results showed that project schools made changes to pupil organisation and teaching from 1993 to 1995. At the classroom level one third had changed

from mixed ability to some form of streaming; one half had moved away from integrating subjects into cross-curricular topics towards subject based teaching; one quarter had decided to do more whole-class teaching; just under a half had introduced regular formal testing in years 3, 4, 5, and 6.

At the school level, almost three quarters of the schools had discussed the 1995 results fully as a staff and this resulted in policy changes, affecting all classes; just under a half of the NAPS schools had made alterations to their approaches to the marking of writing and to teaching of reading and spelling; one third had changed their mathematics policy in order to present mathematics in the way that children are 'far more likely to meet on a test paper'; a small minority of schools had introduced revision time for science during years 3, 4, 5.

At the teacher level teachers said: they were using the results to inform future teaching, for curriculum planning and for the review of assessment practices; their preparation before the testing began (practised timed tests, familiarising children with test style and layout, setting mock tests and revising material), had helped the children get good scores; they had introduced an assessment format that matched the SAT ones so that the children were more familiar with it and the assessment tasks were changed to be of the same time and length as the national testing ones. Curriculum content was included if it would be in the testing. Teaching styles (e.g. dictation) mirrored the style of the national testing. More time was given by the teachers to revision, practice tests, timed work, tests at the end of a topic.

In summary, the national testing programme had an impact on schools: results have been used to identify aspects of teaching which heads and teachers perceived as influencing pupils' results; the majority of teachers made changes to their teaching (even when the scores were the same as their own assessments). The teachers used the results to identify how they might best prepare the next group of children for the SATs.

A significant comment came in the analysis of the use of the information gathered by secondary schools. Twenty of 21 secondary schools had received the national testing information from feeder primary schools (although not always from all feeder schools). In two schools, the results were used to stream students. The results were not used for curriculum planning in the other 18 schools.

The report sought information from parents. A significant number of parents in the first year (1995) could not interpret the results despite efforts by schools to explain them through a variety of means. The authors note that if a significant number of parents do not understand the results, it can be reasonably assumed that some parents would not be able to interpret league tables of school-by-school results.

Harlen, Malcolm and Bryne (1995) gathered evidence during a four year evaluation of the national curriculum and assessment implementation. They sent surveys to 200 primary schools and undertook interviews in 25 of the schools every year for four years. The evidence from the evaluation suggests that the changes taking place are such that assessment is being included in lesson planning more than before; being carried out in the classroom more consciously; encompassing a wider range of learning outcomes; taking place within a school-wide framework; more often than before recorded in a structured manner.

Their analysis is that there were definite moves towards making teachers' assessment more purposeful, more wide ranging, more systematic and better recorded. They note that there is not much evidence of the extent to which this was used to inform teaching decisions; nor how continuous assessment was summarised to give a summative judgment. In these authors view it was important to have banks of tasks which involved sharing of different teachers interpretations if there was to be a close link between teacher assessment and pedagogy.

Mehrens (1997) in a Presidential address to the American Education Research Association concludes that high stakes national assessment will result in higher test scores. However, both test security (and insecurity) and the opportunity to mis-administer and mis-score tests must be considered in evaluating whether increased test scores represent increased knowledge and learning with reference to the mandated curriculum. Mehrens also states that there is not enough research

done to date on the consequences of assessment, and that existing research findings cannot indicate if the consequences of assessment are positive or negative. The decision about whether a consequence is positive or negative is related to differences in convictions about the proper goals of education. Mehrens maintains that depending on your values and convictions about education, there are both positive or negative consequences of assessment with respect to:

- curricular and/or instructional reform,
- increasing teacher motivation and stress,
- changing students' motivation or self concepts,
- decreased scores on assessment, and
- public awareness of student achievement.

Three important New Zealand research studies have recently provided information about assessment-learning links. First, Bell and Cowie (1997) reports a detailed study of ten teachers using formative assessment. They report positive gains in understanding of pedagogy as the teachers implemented processes and interacted with the researchers. Teachers' planning of lessons, finding out if the learning activities were working, finding out if the students were learning the intended learning outcomes, monitoring the student and their own progress, and planning for future units of work all benefited from the use of formative assessment.

Second, Moreland and Jones (1999) showed significant gains for student learning when formative assessment and teacher subject knowledge are enhanced.

Third, Keown (1996) reports research involved eight secondary school departments which were trialing New Zealand Unit standards. This involved 16 teachers and some 80 students and took place over a six month period. The project was designed firstly to ascertain how teachers gathered evidence and made decisions on the process of collecting the quality and quantity of evidence required to convince an assessor that a candidate is, or is not, competent in relation to the function defined by the element without over assessing or under assessing. The project noted a gap between the standard expected by central authorities and those expected by the school. Keown notes that to close this gap schools need training programmes which will assist teachers in expanding their repertoire of assessment strategies, and that there is a need for more opportunity for dialogue and discussion between schools and within schools. A clear policy of teacher development in assessment skills is further recommended.

In a recent major Education Review Office report on 180 New Zealand schools the summary concluded that there were two clear messages from the review of this large sample of schools. First, teachers need help with their assessment strategies if they are to do an adequate job in all primary school curriculum areas. Each school should not have to devise its own approaches. There should be national procedures available to schools. Second, teachers would need considerable in-service education to learn how to understand and use these national procedures. In the light of our literature review there is a problem with advice leading to what would become high stakes assessment procedures. Resources would more likely contribute to enhanced learning if appropriate formative assessment processes were encouraged and teacher development programmes consistent with them were introduced.

Summary

This section complements and expands on the analysis in Chapter 4, Section 4.2.3, providing further detail of the interactions in the analytical framework for the review (Figure 2). The review finds that many external assessment processes have an inhibiting effect on pedagogy, focusing teachers' interactions with students towards retention and recall of factual material rather than the wider aims of the curriculum. The major potential advantages for student learning from formative assessment would be hampered by this restricted pedagogy.

5.5 Effects of assessment on learning and social outcomes

The discussion so far in this chapter has continued to elaborate a complex interaction between curriculum, assessment and pedagogy, which through the intervening role of the teacher comes to affect students' learning and social outcomes (see Figure 2). There is research evidence for formative assessment having the potential for significant effects on student learning and social outcomes. Discussion of this topic will be deferred to Chapter 7 after consideration of pedagogy in the next chapter.

General Summary

The research reviewed indicates that:

- high stakes national testing can create distortions of the mandated national curriculum objectives and produce undesirable classroom effects.
- there is a positive relationship between increased formative assessment and increased learning outcomes.
- summative assessment can dominate in classrooms, resulting in fewer opportunities for formative assessment.
- there is a strong effect by assessment approaches on pedagogy, both on teaching activity and on content taught. In particular, high stakes national testing and monitoring strongly affect pedagogy and can create distortions of the mandated curriculum objectives when memorisation of facts and displays of basic skills was promoted, rather than higher levels of learning and skills which can be assessed through performance tasks.
- in high stakes national assessment, there was evidence of 'teaching to the test'. This resulted in a restricted pedagogy and curriculum for students; changes to student organisation in the classroom and school; a move away from integrated studies to subject-based teaching; a move towards more whole-class teaching; a move towards teaching styles mirroring the style of national testing; and a move towards content similar to that in the testing.
- assessment in the classroom was seen as being more systematic, more purposeful, better recorded, and was discussed more by the teachers, when lesson or unit planning involved more consideration of the assessment of learning outcomes. The effect was also seen as positive if the learning outcomes matched teachers' convictions and goals of education. Exemplars of good practices need to be developed consistent with appropriate learning outcomes.
- there is a growing body of research in classrooms indicating the advantages and disadvantages of different forms of assessment tasks on student learning outcomes.
- concern has been expressed about the quality and usefulness of data from large scale summative assessments (for example, IEA, TIMMS) in the development of national assessment policy. Furthermore, large scale studies like these assessment procedures are only indirectly connected to what teachers actually teach. The relationships between curriculum and teaching approaches, and student achievement are difficult to establish.
- there were significant research gaps. There is little if any research (particularly in New Zealand) on whether national testing and other accountability measures promote education change or development. No research findings on assessment activities which would provide valid, reliable, trustworthy and fair information about Maori, Pacific Nation or other minority students was found in the review.

Chapter 6: Pedagogy

Introduction

A considerable number of aspects of teaching and learning are subsumed under 'pedagogy'. In reviewing the literature in this field, it became clear that the teacher was being identified as playing a pivotal role in the relevant research. The intervening role of the teacher (see Chapter 3, Section 3.2) is most obvious in this literature and forms the focus of much of the research. For this reason, Chapter 6 is structured so that the role of the teacher, and as a consequence teacher development, is explored to provide a focus of the papers reviewed. Section 6.1 presents the research literature on pedagogy. The literature in this section is organised around four major foci: what makes a good teacher (6.1.1), teacher knowledge of the subject (6.1.2), different groups in the classroom (6.1.3), and class size and organisation (6.1.4). As the position of the teacher in the teaching and learning process is central to this chapter, Section 6.2 provides a more detailed commentary on the nature of this role in the wider context of systemic educational reform. Included in this section is a commentary on teacher professionalism. Since the central role of the teacher raises the question of professional development to support change, Section 6.3 focuses on the literature that researches the professional development of teachers. A brief comment is then made in Section 6.4 on the effects of pedagogy on student learning and social outcomes.

6.1 Research on pedagogy

The appropriate entry into research on pedagogy, since in many guises it has been the research focus, is to seek answers to the question of what makes a good teacher. This question is explored in Section 6.1.1. In Section 6.1.2 the focus is teacher knowledge of the subject. In Section 6.1.3 the research focus shifts to whether there are different pedagogical strategies appropriate for different groups in the classroom. The effects of classroom organisation and size on pedagogy is the research focus in Section 6.1.4.

6.1.1 What makes a good teacher?

Although it is beyond the scope of this review to include research beyond the past five years, this poses a particular difficulty for research on teaching because much of the seminal research on teaching was conducted outside this period. Since the advent of structured empirical research on teaching in the 1950s there has been a consistent interest in the way teachers affect the achievement of students. Research in the 1950s and 1960s was generally experimental and was often carried out in laboratories as it was assumed that there was a straightforward relationship between teaching and learning, and that through this research, a science of teaching could be identified. The equivocity of these research results meant that from the 1970s, researchers interested in how teacher behaviour affected student achievement began to gather data by observing and coding classroom activity (Evertson, and Green 1986).

Although there are some exceptions, the vast majority of studies on teacher behaviour and student achievement have been quantitative in nature and while researchers continue to search for links between teacher behaviour and student achievement, a large proportion of the relevant body of research was conducted in the 1970s and 1980s. Thus, findings of Brophy and Good's (1986) extensive meta-analysis of 204 process-product studies relating to teacher behaviour and student achievement is summarised.

From the outset of this review Brophy and Good (1986) make it clear that the effect of teachers on students is quite different from "teacher effectiveness" because the latter includes other aspects of the educative process such as the social and moral development of students. Brophy and Good's analysis (1986) is confined to reviewing classroom-based observational studies which investigate the relationship between teacher behaviour and students achievement in terms of mastery of the

curriculum. Synthesis of the studies revealed that student achievement was maximised when teachers:

- provide opportunities for students to master lesson content;
- are task-oriented and emphasise academic achievement as an important goal;
- use effective routines so that activities run smoothly, and time spent on organisational or management is minimised;
- implement activities that match student achievement levels and needs;
- maximise coverage in such a way that activities are paced so that students are able to learn each objective and integrate this knowledge with past learning;
- maximise the time they are actively teaching students.

Although there was a paucity of data that explained the effects of whole class versus small group teaching in the research reviewed by Brophy and Good, there was a positive correlation between teacher led instruction and student achievement, and a negative correlation for when the students were engaged in individual work without direct teacher supervision (Brophy and Good, 1986). Moreover, aggregation of the results of the 204 studies showed that the way teachers structure lessons and communicate with students has a considerable effect on students achievement. In summary, student achievement is maximised when teachers:

- carefully structure activities so that students are presented with features such as advance organisers, links, analogies, overviews, outlines and reviews;
- repeat main points regularly;
- questioning and presentation styles are characterised by clarity;
- are enthusiastic;
- match the level of questioning with the activity (closed questions which the students can easily answer for rote learning activities, and open-ended questions that engage students in high-level cognitive processing when teaching a complex cognitive content);
- use wait time after asking a question;
- call upon a variety of students to answer questions;
- provide positive but accurate feedback about student replies to questions;
- incorporate students' comments and questions into the lesson.

Brophy and Good (1986) warn that although these findings can be generalised to all school instructional settings, some aspects of the summary are context specific. The age level of the class was a contextual characteristic that had an effect on classroom management styles and the way teachers presented activities and interacted with students. Similarly, low-achieving students, or those from low-income backgrounds benefited when the teacher broke activities into smaller structured steps and interacted directly with the students.

In conclusion Brophy and Good (1986) remark that analysis of these process-product studies clearly shows that educational change efforts must take into account the complexity of teaching, rather than trying to solve instructional deficits with simple prescriptions. Moreover, effective teaching relies on many aspects of teacher practice, and is profoundly affected by contextual factors such as the characteristics of the students and the subject being taught (Brophy and Good, 1986). This differentiation is important, because while a teacher's pedagogical practices are linked to their efficacy as teachers, these factors are only a part of the effectiveness picture.

The studies and analyses of research reported in this chapter represent a balance between qualitative and quantitative research methodologies. Qualitative studies provide contextualised and detailed information about how teachers operate in their classrooms (Erickson, 1986) and quantitative methodologies are more effective at isolating the effects of individual teacher behaviours (Linn, 1986). Overall, there has been an effort to avoid studies that rely wholly on teacher report although some of the larger scale self-report studies have been included simply because there is an overall scarcity of recent large scale classroom-based research investigating teacher practices and their effect on students.

Shulman's frequently-cited analysis (Shulman, 1987) of the knowledge bases for teaching details seven categories of knowledge that an effective teacher should possess:

- content knowledge
- general pedagogic knowledge, with special reference to those broad principles and strategies of classroom management and organisation that appear to transcend subject matter
- curriculum knowledge
- pedagogical content knowledge, the professional understanding of the interaction between content and pedagogy
- knowledge of learners and their characteristics
- knowledge of educational contexts, ranging from the workings of groups in the classroom, governance and financing of schools, to the character of communities and cultures
- knowledge of educational ends, purposes and values and their philosophical and historical grounds.

To date it has proved impossible to quantify each of these knowledges in an effective teacher. Shulman argues that teachers are a complex mix of these seven categories. Any discussion of pedagogy must also recognise the complexity of the classroom which is where these knowledges come into action. These knowledge bases can be linked to the search for indicators of teacher effectiveness. Many thousands of studies, particularly in North America, have sought to identify teacher skills and characteristics and relate them to student achievement. A problem in much of this work is that researchers attempted to single out particular variables to measure their effects, and in so doing ignored other interacting variables in the complex teaching-learning settings. The search for a science of teaching in which effectiveness variables were known and could be trained into teachers was not realised. Today there is much more acceptance of teaching as a subjective activity, more an art than a science, which depends on well-prepared teachers using a repertoire of professional skills, abilities and judgments to make classroom decisions (Eisner, 1994).

The analysis above should not be taken to suggest that little is known about effective teaching. Porter and Brophy (1988) summarised a large number of correlational research studies on effectiveness factors, and reported 11 characteristics. An effective teacher:

- is thoughtful about practice and reflects upon it,
- develops clear curriculum aims and objectives,
- makes clear expectations of students,
- knows the subject content,
- knows the students' characteristics and needs,
- uses a variety of sources to enrich learning,
- uses a variety of objectives,
- teaches students how to employ meta-cognitive learning strategies,
- integrates subject matter between subjects,
- provides regular feedback to students about their learning,
- accepts responsibility for student achievement.

Another important observation about teacher effectiveness is that effectiveness must be matched with purpose. That is, particular teaching styles are designed to achieve particular student outcomes. For example, Joyce, Weil and Showers (1992) reported the use of an inductive model of teaching where a team of researchers developed inductive questioning and discussion as a technique to improve students' learning achievement about social issues in social studies. This design had a purpose. Other educational objectives require other models of teaching. A single model appropriate to all educational objectives is an illusory goal, unless it is so complex as to be unworkable.

Porter and Brophy's list of characteristics links with the research of Schon (1987) on teachers as reflective practitioners, Eisner (1994) on connoisseurship, and Hoyle and John (1995) on the teacher as an extended professional. In all these complementary descriptions the teacher is much more than a narrowly-trained technician. A New Zealand focus on this issue came from Haigh and Katterns (1984) when they described an effective teacher as able to:

- draw upon knowledge from a wide repertoire of teaching modes and use a variety of skills to create learning situations that suit the students and their intentions;
- use a wide repertoire of teaching strategies and tactical skills to match the intentions for student learning, either by ensuring that the plans are put into practice and followed, or by responsive moves;
- be sensitive to cues and signals in the teaching-learning situation so that they sustain or modify their planned actions.

Gipps (1999) in a major review of primary teaching and learning has a similar analysis of effective pedagogy, and documents the actualities of teachers responding to innovations. Research data were gathered from year 6 teachers in the UK on the way they teach and how this related to learning and the impact of the national curriculum and Assessment programme on their teaching. Four LEAs were involved, 8 schools in each LEA, one teacher per school and the Head, a total of 32 teachers and 32 Heads. In five schools, chosen so that each was significantly different, extensive case studies involved classroom observation and teacher interviews. The authors concluded that learning and teaching was highly complex and could not be subsumed under any particular model. They suggested that most primary school teachers adopted an eclectic approach, but that frequently they were confused about which approach might be most appropriately used in any particular context.

Gipps comments that we need to talk in terms of teachers having not a pedagogy but a range of pedagogical strategies (which includes a range of materials and content, teaching style, and classroom rules/agreements) to be used at different times. The author's view is that assessment to support learning, formative assessment, must be an integral part of these pedagogical strategies because teachers must reflect on the pupils' understanding: the pupils' understanding is a key element in choosing the appropriate pedagogical strategy.

External pressures on effective teaching are indicated when, with regard to the National Curriculum, while not one teacher believed that children learn solely from the transmission of facts, one quarter of the teachers had felt pushed into more whole class teaching of material because of the amount of work to be covered by the National Curriculum and associated assessment processes. This constrained the teachers' use of children's ideas as starting points for work and they regretted this.

Recent empirical research provides further evidence about factors that relate to achievement. In a large-scale research project, Wragg, Wragg, Hayes and Chamberlin (1998) looked at what school principals do to raise literacy achievement in primary students in the UK, what happens in the classrooms of successful teachers, and which individual children improve and why they do so. Questionnaires were completed in nearly 1,400 schools, principals and LEA personnel were interviewed in four local authorities, 35 teachers' classrooms were closely studied over a year up to Year 6 level, and intensive observations made of 258 pupils to find which of them improved and why they did. The authors acknowledged the difficulties in measuring improvement and attributing this to particular causes, noting that classes of children vary, sometimes considerably, in their backgrounds and abilities. One way of measuring improvement was by test scores, but other factors needed to be considered, such as (in this case) attitude towards reading, and breadth of reading material. Improvement was measured by quantitative and qualitative outcomes from interviews, questionnaires and observations, and observations of lessons, and achievement tests. From the considerable amount of information gathered it was concluded that single causal factors could not easily be identified. As Wragg et al point out, "classroom life teems with thousands of micro-episodes, often hanging together in strings and clusters" (p.258). Surrounding them are thousands more outside the classroom that influence schooling, sometimes reinforcing, enhancing,

or undermining. Improvements in reading may result from some combination of these influences, inside and outside the classroom. A particular cause of improvement may have only a minimal effect. In the 35 case study classrooms it was not unusual to have some pupils in the same class improving while others declined. Perhaps not surprisingly, it was difficult for the researchers to extract "universal truths" from the classroom observations. Even so, ten features of classroom pedagogy seemed to be linked to success. These were: teacher enthusiasm for literature and use of varied resources; a sound knowledge of books and teaching strategies, making literacy important; celebrating good pupil progress publicly to build confidence; matching reading interest to particular children and resources; systematic monitoring and assessment; regular and varied learning activities; development of pupil independence to take their own reading forward; sound classroom management; effective pupil-teacher relationships and high levels of on-task behaviour; and high positive expectations, no matter what the pupil level. There is a good deal of consistency with the factors already outlined by Porter and Brophy and others in this section.

The report of Wragg et al (1998) discusses pupil autonomy as a complex issue important for teachers. The teachers in their study had difficulty finding a correct balance. On the one hand giving too much responsibility to pupils could amount to not knowing whether understanding had been achieved, or whether the pupil was selecting reading material wisely. On the other hand, too much checking could mean time wasted in queries, over-dependence on the teacher and a lack of pupil confidence in managing at least some of their own reading. In this study some children surprised with their perseverance and determination to improve, as did some teachers and parents in their efforts to help particular children.

Teaching students to generate questions has been identified as a pedagogical technique that assists students to develop comprehension of text they read. In an attempt to review the effectiveness of this teaching approach, Rosenshine, Meister and Chapman (1996) reviewed 26 classroom based intervention studies in which students were taught this technique and their comprehension measured at the end of the intervention. In 17 of these studies students were taught the technique via regular instruction while in the remaining 9 studies the technique was taught through reciprocal teaching. The method of reciprocal teaching involves the teacher modelling the approach and then providing gradually decreasing assistance as the students attempt the task. A synthesis of the results of these intervention studies clearly showed that teaching students to generate questions as they read text resulted in gains in comprehension. There was no significant difference between the results of students taught the approach through reciprocal teaching or traditional teaching methods. However, the researchers found that the type of procedural prompt used by the teachers involved in each study also had an effect on student comprehension. The most effective prompts such as using question stems and signal words were also the simplest, suggesting that students make greater gains in comprehension when they can quickly adapt to the approach. Rosenshine et al (1996) conclude that future reviews of studies in other areas of cognitive strategy should examine the effects by grouping them in terms of procedural prompt to facilitate an understanding of why some studies are successful and other are not.

There are many other research reports complementary to those outlined above.

Harlen and James (1996) describe how student learning experiences should be well matched to the existing point of development of their ideas, skills, attitudes and values; have continuity with, and build on, previous experience; relate to current interests and experience; be perceived by learners as relevant, important, stimulating and valued for themselves, rather than simply for their usefulness in passing tests and examinations. They describe the need for the teacher to: have a thorough and deep understanding themselves of the subject matter to be taught, how students are likely to learn it and the difficulties and misunderstandings they are likely to encounter; have a clear idea of the progression in the ideas, skills etc. which are the goals of learning and the course students are likely to take in this development; be able to recognise the point in this development reached by their students; know and be able to use various strategies to find out and to develop students' ideas skills, etc. Harlen and James comment that teachers cannot provide experiences and

activities that guide students' progress towards understanding of ideas if they themselves do not know what these ideas are; nor can they provide experiences that challenge students' misunderstandings if they themselves share the same misunderstandings. Although teachers generally have sound pedagogic skills to carry them through 'difficult' aspects of the subjects they teach, the result for students are restricted and impoverished learning opportunities, particularly for learning aimed at understanding. For example, recent research into teaching science at the primary level (Harlen, Holroyd and Byrne 1995) indicated that the coping strategies adopted by teachers with poor background knowledge included relying heavily on prescriptive work cards or texts giving students step-by-step instructions, emphasising process aims and neglecting conceptual development, minimising students' opportunities to ask questions, avoiding practical work and any equipment likely to go wrong.

A parallel view comes also from the substantial research programme in the US reported by Newmann and colleagues which focused on 'authentic pedagogy' (which they describe as: using a prior knowledge base from one or more fields; striving for in depth understanding rather than superficial awareness - which means that learning is concerned with central ideas, information is linked, a few ideas are covered in depth, knowledge is constructed and based on problem solving, information is evaluated, and learning can be re-applied; expressing conclusions through elaborative communication which makes connections between substantive knowledge and either public or personal experiences). In authentic pedagogy pupils are given props to assist their learning, and are challenged to take their learning beyond reliance on their own constructions of knowledge.

Newmann, Marks and Gamoran (1996) report research which studied mathematics and social studies classes in 24 restructured schools (described as highly innovative) over one year. The authors provide empirical evidence that authentic pedagogy pays off in student performance across different grades and subjects. Significant findings for pedagogy were that:

- while some classrooms in schools were far more successful than others, the overall levels of authentic pedagogy observed fell well below the highest level of the proposed standards.
- authentic pedagogy resulted in improved academic performance for students at all grade levels in both mathematics and social studies.
- it is possible for schools to provide authentic instruction reasonably equitably and its effects on students' academic achievement is reasonably equitable. In the sample of schools used in the study, neither gender, race, ethnicity, nor social economic status were considered to determine which students received high or low levels of pedagogy. The researchers claim that none of these influences affected ethnic academic achievement.

The design of the study leaves open the question of whether a clear causal relationship is established between pedagogy and learning, but the relationship between authentic pedagogy and student performance suggests reasonable grounds for these claims. A further comment would be that even in supportive schools the task of improving pedagogy was found to be enormously difficult.

In a more elaborated commentary Newmann and Wehlage (1998) synthesise 5 years of research conducted by the Centre on Organisation and Restructuring of Schools (CORS). The research involved several thousand students in hundreds of schools. A conclusion was that the recent education reform movement gave too much attention to changes in school organisation that do not directly address the quality of student learning. A further conclusion was that student learning can meet high standards if educators and the public give students three kinds of support - teachers who practise authentic pedagogy, schools that strengthen professional community, and supportive external agencies and parents. This commentary emphasises the value of independent work structures and staff development for teachers that might result in enhanced student achievement.

There is considerable complementarity between the comments of Newmann and Wehlage above and the research report by Levine and Lezotte (1995) which attempts to identify the characteristics

of "effective schools". Their specifications include: a productive school climate and culture reflecting shared values; a safe and orderly environment; teacher commitment to a shared and articulated mission focused on improving achievement; a problem solving orientation; staff cohesion, collaboration, communication and collegiality; staff input in decision-making; a school-wide emphasis on recognising positive performance.

A similar difficulty in the uptake of authentic pedagogy reported by Newmann et al (1996) comes from Roelofs and Terwel (1997) in their study of the introduction of a new national curriculum for 12-15 year olds and related pedagogy in the Netherlands. The curriculum required that more attention be paid to the links between subjects, and for teachers to incorporate the characteristics of authentic pedagogy in their teaching. Three projects looked at intervention studies in sciences (including mathematics, physics, chemistry and biology) and language education. The number of students in the projects ranged from 400 - 800. The research indicated that authentic pedagogy was not found to any "satisfactory extent" in the first stage of this reform. There was a mismatch between teachers' and their students' views as to how often and how well teachers practised authentic pedagogy. The researchers conclude that there was a possible mismatch in understanding the meaning of 'relevant' and what constituted an 'activity'. In addition, schools did not always fulfil organisational conditions to facilitate authentic pedagogy. Teaching resources and text books, a strong feature of Dutch classrooms, were also not geared to facilitate it, media (e.g. email, CD roms, etc.) were not sufficiently available in schools, and contacts with external institutions were difficult to realise in practice.

Summary

The research in this section indicates that there are effectiveness factors of teaching. These factors include teachers: having a broad understanding of curriculum aims and objectives; having a wide range of pedagogical strategies; having high expectations of all students; knowing their students well; providing good feedback to students; recognising student success; having sound content knowledge of their subject; and having a sense of responsibility for student progress. The literature suggests that the mixture of factors is complex and that they often come together through judicious decision making in the classroom. However, researchers concur that it is extremely difficult to quantify 'a good teacher' and the literature does not support any direct link to student achievement. The literature reviewed in this section does not identify the characteristics of a good teacher of different groups of students, for example Maori students, beyond broad support for teachers knowing their students. This topic is further considered in Section 6.1.3.

6.1.2 Teacher knowledge of the subject

A recurring theme in this review of pedagogy and the effects on learning has been the developing recognition of the importance of teacher knowledge of the discipline being taught. This is increasingly seen as vital. One commentator (Phillips, 1997) even considered that there was conflict between a focus on knowledge and a focus on student emotional and social needs. The other research described below sees a more complementary process, in which acquisition of content knowledge and other attributes such as confidence and meta-cognition work together.

Teacher knowledge of the discipline is related to the use of various assessment processes, since it is crucial that a teacher has knowledge of the discipline to provide direction to learning. When teachers are unsure of their discipline's structure they are not well equipped to guide learning in it or assess that learning.

Corcoran and Goertz (1995) in a major literature review of organisational effectiveness and job performance of teachers suggest three general sets of variables:

- intellectual ability, knowledge and skills of teachers;
- the quality and quantity of the resources available, instructional time and class sizes;
- the social organisation of instruction.

They conclude that research has found positive evidence that content-relative knowledge had a positive impact on student performance. They report other studies which have shown that teachers often lack the subject matter knowledge needed to make the pedagogical changes envisioned by reformers. Related studies show that policy makers often proceed with reforms without understanding what teachers need to know to be able to implement them and that an absence of adequate resources hinders pedagogical reform. These authors comment that our knowledge of what works in classrooms is very limited in terms of research findings.

Teacher confidence is an important factor that shows up in some research. Harlen (1999) in a major survey of research into primary and secondary science teaching, reports a well established finding in many countries that attests to teachers' low levels of confidence about teaching science, and understanding of science concepts. The impact of teachers' confidence and understanding on pupils' learning opportunities has also been established by research, low levels of confidence and understanding being associated with restricted classroom practices inhibiting creativity and questioning. Presumably, these findings would apply to other school subjects, too.

Confidence was again a major factor in another science project. Scott, (1997) suggested a decline in teaching primary science was due to primary teachers of science not being confident in their scientific knowledge and in teaching science. The implementation of the primary science curriculum is affected, though it is difficult to determine the exact nature of the implemented curriculum. In the primary classrooms studied opportunities to learn science were found more frequently in activities not labelled science. Overall the classrooms were dominated by activities labelled language, reading, mathematics and physical education; though Scott notes that many of these activities offered opportunities to learn science

Osborne and Simon (1996) similarly report ample evidence that in England and Wales primary teachers' subject knowledge of science is generally weak; significant gaps in conceptual understanding having been found. In their view more research into this feature of primary education is required since teachers face pedagogical difficulties in teaching material when their understanding is inadequate. They found that the new curriculum, which resulted in the introduction of new topics uniformly throughout the country, lacked sufficient teacher development. This led to many pragmatic responses such as 'getting' by and 'muddling through'. In their view this implementation of a prescriptive curriculum for primary science by teachers lacking essential subject knowledge had the risk of decreasing interest and motivation in science for students. Without appropriate teacher subject knowledge no curriculum, however prescriptive, can achieve uniformity and consistency of experience.

Dunkin and Welch (1996) report on the Australian Teacher Knowledge into Practice research programme, an empirical study, run by the University of New South Wales. The authors conclude that when teachers are unfamiliar with course material or when they are asked to implement material they are unfamiliar with, they tend to adopt more prescriptive, teacher-directed methods.

Turner-Bisset (1999) explored how teachers can be prepared with the wide range of knowledge that they require to be expert teachers. The controversy in the United Kingdom about a specified set of teaching skills being defined for teachers, as against a broader knowledge basis is outlined. As a response the author developed a model of the expert teacher. Teaching is viewed as a complex intellectual and practical activity.

Sternberg and Horvath (1995) also explored what it means to be an expert teacher. In this conjectural, as opposed to empirical research, they concluded that experts bring knowledge to bear more effectively on problems within their domain of expertise than do novices. The expert teacher is knowledgeable, having extensive, accessible knowledge that is organised for use in teaching. The expert teacher is efficient in solving problems within the domain of teaching; able to identify information that is promising with respect to problem solving.

Jones and Moreland (1999) in New Zealand demonstrate gains in students' learning in technology when teachers have a clear subject knowledge base to enhance their interactions with students.

Hill and Hawk (1998) report a research project, Achievement in Multicultural High Schools (AIMHI) being carried out by Massey University. In this evaluative study, the authors identified key teacher qualities that enhance learning and achievement. Key teacher attributes include qualifications and subject knowledge.

Snow-Renner (1998) documents teacher reports about assessment practices in mathematics classroom relative to student opportunities to learn. Survey responses were analysed from 1067 teachers. Overall findings were that students in different classrooms experienced differential opportunities to learn relative to reform-orientated assessments, and that teachers had varying capabilities for implementing these assessment practices. Variations were partially attributed to teacher capacity and knowledge.

Sizmur and Sainsbury (1997) background the development of criterion referencing and its role in National Curriculum Assessment. In this discussion the authors argue that teachers' professional knowledge is central to the process of teacher assessment. Therefore, teachers need to have a thorough understanding of the curriculum subject matter, how children learn and how these interact, and that these matters are central to making criterion referencing work in the new curriculum.

Phillips (1997) argues that high expectations for student achievement, clear achievement oriented goals, the amount of time spent on instruction, and the amount of homework assigned to students has a strong relationship with academic achievement. He further comments that maintaining effective relations with students rather than imparting skills and knowledge ('liberal paternalism') motivates teachers to focus on students social and emotional needs to the neglect of their academic needs. He provides a model of school effectiveness described as placing academic learning at its centre, rather than catering for students social and emotional needs as the primary goal.

Gipps, McCallum and Brown (1999) provide a more detailed analysis of the factor of teacher subject knowledge and its impact on learning. Their discussion of the importance of teacher knowledge and teacher expectation of their students is based on the findings of research into primary classroom teaching and learning. Key points are:

- low teacher expectations produce low motivation to learn and erode the academic self-image of students. This leads to under-achievement.
- optimal conditions for learning include a positive, personal learning attitude where challenge is high, and anxiety and self-doubt is low; praise and encouragement; high levels of expectations of all children; high levels of work-related talk and discussion.

The teacher -learner relationship is crucial to the learning process. This includes the nature and purpose of questioning; the need for the teacher to understand the learners' view; teacher feedback as part of formative assessment.

Highly effective teachers of mathematics thought that being numerate requires a rich network of connections between different mathematical ideas, so they used teaching approaches which connected different areas of mathematics; encouraged pupils to describe their methods and their reasoning; and used these for developing understanding and establishing connections. These teachers ensured that all pupils were being challenged and stretched. Similarly, highly effective teachers of literacy generally placed a high priority on meaning in their teaching of literacy, and taught the conventions of reading and writing in a systematic way. They also communicated the functions of language and structured their teaching materials and approaches. Their summary is that teachers need to have a good knowledge of their subject as well as classroom management skills, and understanding of how students learn.

In 1990-1991 the International Education Association (IEA) conducted a literacy survey in 32 countries (Elley, 1994). Data was gathered from classes or grades in which the majority of students were nine years old and fourteen years old. In each participant country the sample of students from each of these levels numbered close to 800. Comparison of the literacy results between the countries indicated that there was a positive correlation between the level of student literacy and the educational achievement of the teaching force. New Zealand also participated in the Third International Mathematics and Science Study (TIMSS). The study involved gathering data about the maths and science achievement of 4935 nine-year old students. The results of this study showed no correlation between the general qualifications of the teachers and student achievement in science and mathematics (Garden, 1996). A further part of TIMSS was the study of the science achievement of 6911 form one and three students (aged 14 years old). The results showed that the students who had teachers with advanced qualifications did achieve higher results in science. This positive correlation was stronger for the form three students than it was for the form two students. (Garden, 1997).

Alexander, Wilcocks and Nelson (1995) report findings from the CICADA Project, (Changes in Curriculum-Associated Discourse and Pedagogy in the Primary School) in England and Wales. Their data were collected through teacher interviews (12) and classroom observations (60 sessions), and a questionnaire survey of a national sample of primary teachers (700 sampled, 536 responded - 77% return rate) as part of PRINDEP (Primary Needs Independent Evaluation Project). In assessing the impact of the National Curriculum the authors arrive at 3 propositions:

- that the National Curriculum has had a major impact on teachers
- that the impact on students is debatable
- that the significant ingredient in this discrepancy is pedagogy.

A major finding was that there was a relative continuity in pedagogy described as a resilience at "the deeper levels of pedagogy" (p111). They continue by saying that any study on the impact of the National Curriculum needs to chart and compare the versions of knowledge, understanding and skill which teachers create for students, and which students create for themselves out of the learning tasks teachers provide.

The authors call for further research into the effect of pedagogy on student outcomes, and the relationship between subject structures, the teachers' conceptual grasp of the subject matter, and the character and content of teacher-pupil interaction.

Gitlin and Margonis (1995) contribute further to the discussion of teacher knowledge when describing how this should contribute to curriculum development, and how if this does not occur, and the curriculum has an uncomfortable fit with teacher professional knowledge, serious difficulties are likely. Teachers are liable to resist change in these circumstances. Teachers' resistance to change may, however, be positive, in that teachers' good sense and professional knowledge may mean that they evaluate the impact of their own changed practices on students and student achievement.

Morris (1995) provides a student perspective on teacher knowledge. This research explored the in-school and out-of-school factors which promote or inhibit the learning of senior boys and girls in the secondary school setting. Questionnaire responses from over 400 students identified the best teachers as being skilled communicators, who knew their subject and were personally interested in each student.

Summary

This section argues that teacher knowledge of their subject (content knowledge), how students learn (pedagogical content knowledge), and the interaction of these two factors are essential to support learning. Good teacher knowledge of the content of their subject was found to have a positive effect on decisions to change pedagogical strategies to create learning opportunities. In addition, sound content knowledge seems to have a positive effect on planning, assessment,

implementation of curriculum and curriculum development because the teachers are more able to make necessary pedagogical changes. The literature reviewed also suggests that low level content knowledge is associated with restricted classroom practices which impact on student learning opportunities and on teacher confidence in teaching the subject. This is particularly so for primary teachers. However, there remains doubt about the level and extent of teacher knowledge that demonstrates good and unsatisfactory teacher knowledge.

6.1.3 Different groups in the classroom

Pedagogy is the area that has provided the most research associated with different groups in the classroom. The notion of 'appropriate' pedagogy for 'minority' students is frequently based on the belief that the traditional didactic approach to teaching is not suited to their learning, since they may have different ways of learning and differing responses to teachers, to other students and to the subject itself. It is generally argued that a variety of pedagogical strategies are what is needed to interest the more marginalised groups. Part of this approach is also concerned with providing classroom settings that are culturally comfortable for different groups of students.

In many of the studies reported above students have been described in a generalised manner. The complexity of the classroom discussed by Wragg et al (1998) requires the pedagogical focus to shift to the individual learner. The most frequent research response to this need has been to group individuals into structured groups, using culture, gender and other organising characteristics. A significant problem with this approach is that even in seemingly similar groups there may be marked variations among the individuals. For example, in the New Zealand context it would be unwise to treat all Maori students as if they had the same background experiences and the same prior knowledge and skills.

This section documents research which looks at different groups in the classroom, frequently making generalisations which would need to be carefully implemented for individual students.

Mehan, Akamoto, Lintz and Wills (1995) provide a commentary on ethnographic studies of multicultural education in the classroom and school where the emphasis was on attempts to be culturally inclusive by modifying classroom organisation, discourse patterns or the curriculum. The authors reviewed studies primarily carried out in bicultural or monocultural classes, and came to five generalisations:

- Instruction should be academically rigorous (maintaining high expectations, seeking comprehension rather than just decoding, and sense-making instead of decontextualised skills) accompanied by a system of social support.
- Classroom and discourse organisation should be student-centred (small group instruction, cooperative learning, facilitating student-initiated participation) so that students are engaged and perform at a higher cognitive level.
- The teacher should become an ethnographer - instead of teachers finding out about the generalities of the ethnic traditions of different cultural groups in the abstract, they should learn about the details of their students' lives in the particular, that is gain context-specific cultural knowledge.
- Students' knowledge and experiences should be a resource for engagement and productivity.
- General principles should be adapted to local circumstances.

Caution is needed with these generalisations, for the Mehan et al study (1995) has only weak empirical support for its claims of causality.

Several other studies explore the cultural dimension of pedagogy.

Irvine and York (1995) provide a critical review of the research on the learning styles of culturally diverse students. Their review suggests that the widely held claims that African American,

Hispanic and Indian students are field-dependent learners who prosper academically when taught with field-dependent teaching strategies are premature and conjectural. However, it does suggest that learning styles research has significant potential for enhancing the achievement of culturally diverse students when coupled with exploration of teachers' actions, instructional goals, methods and resource materials in reference to their students' cultural experiences and preferred learning environments.

Culturally-responsive teaching strategies are explored by Phuntsog (1998). This paper reports a literature review of 13 documents published between 1992 and 1997, and profiles crucial aspects of a culturally-responsive pedagogy. The author proposes a holistic framework for integrating different levels of culture into culturally-responsive teaching. The research suggested that culturally-responsive teaching encompasses respect for diversity; creation of a safe, inclusive, respectful environment; integration of responsive teaching practices in all disciplines; and transformation of curriculum to promote social justice and equity in society. The literature review highlighted problems and issues for developing culturally-responsive teaching. A main finding was that educators' attitudes play a vital role in either empowering or disabling diverse students. The proposed framework integrates salient features of culturally-responsive practices, recognising the teacher's central, crucial role. It shows interrelationships between four levels of culture (personal, microculture, macroculture, and global culture). The paper identifies 5 essential features: cultural literacy, self reflective analysis, caring and inclusive classrooms, respect for diversity, and critical pedagogy/transformational curriculum.

A related New Zealand study (Chapple, Jeffries and Walker, 1997) sought research evidence for suggested reasons for the gap in Maori educational outcomes, focusing on issues of pedagogy, and also wider institutional and social issues. They report a disappointing lack of research knowledge about these gaps. Their suggested reasons for the gaps, and a brief commentary follow (this topic will be returned to in Chapter 8, Recommendations for research).

- Tastes and preferences - i.e. the values of Maori, such as not wanting material goods, including a Maori preference for different institutions using Maori language, pedagogy, content and curriculum. The report found the research to be inconclusive in this area.
- Under-investment - includes discrimination in the market place that reduces their return on acquiring higher skills, and a Maori/class interaction. No research in this area for Maori has been done.
- Family resources - money, inclination, time, quality of input, knowledge of the system. There is considerable evidence that on average Maori families have fewer resources available to ensure Maori children succeed in the education system.
- Peer pressure - including the effects of colonisation, role models, parental encouragement, and social status. One study explores and supports this reason but overall there is a lack of evidence.
- Education supply problems - curriculum and school structure factors. This covers school curriculum, low expectations of Maori by teachers, unconscious teacher bias, role models, organisation of the school, nature and form of assessment, and the Maori community. There is a mixture of research on these topics which needs clarification.

Teacher expectations are commented on above, and known to be important in student learning. In the Strengthening Education in Mangere and Otara Evaluation Project (Timperley et al, 1999, pp 51-52), it was found that 50% of Year Three students in one large primary school were one and a half to two years below national norms on literacy measures. The delay of up to two years continued throughout the children's primary schooling. The teachers focused upon improving literacy skills, and developed 25 skills seen as vital for a child to possess for success in school literacy programmes; skills such as relating print to messages, independent work skills, following instructions, and using various "tools" like scissors.

A survey of 40 new entrants on their mastery of the 25 skills as they entered school showed that, on average, 74% of the skills had been achieved already. These results differed from the teachers' original predictions that only 30-40% would have already been mastered. Just seven children

scored less than 50% and 15 scored over 90%. Interestingly, the one teacher who had expected high levels of mastery on school entry, also achieved the highest reading outcomes with her students. These findings show that, at least in this sample, more children were ready for literacy learning than was predicted by their teachers.

Still in the New Zealand context, Clark (1998) reports on Maori and Pacific Island student participation and performance in mathematics.

Maori

While Maori participation in Mathematics has increased over recent years (although still very low at the higher levels), Maori achievement/performance has not increased significantly. The mean mathematics score for Maori in the latest TIMMS results is significantly lower than the New Zealand mean score, and lower than students on most other countries. Maori girls score lower than Maori boys. There is some indication that students in Kura Kaupapa Maori are doing better than other Maori. Reasons cited include low expectations of teachers/schools, Maori students being overly categorised as non-achievers by teachers/schools, a coinciding of immigrant teachers with predominantly Maori schools, and strong peer pressure.

Clark suggests that to raise teachers' expectations of Maori students:

- teachers should be more persistent and aware. Poor language skills does not mean a lack of understanding,
- teachers should use culturally appropriate environments and contextualised examples,
- teachers should be more informed with facts about Maori students in mathematics.

Pacific Island students

Pacific Island students have substantially lower mean scores in the latest TIMMS results than all other groups in NZ. Clark claims that this is directly related to cultural factors, particularly to teachers' unawareness and misunderstanding of Pacific Island mores. Her proposals for remedying this problem include: more appropriate teaching methods, increased teacher expectations and homework centres. These proposals are based on reported experience rather than research, and the large gap in research evidence is returned to in Chapter 8.

Collins (1994) reports a review of factors to do with compensatory programmes for disadvantaged children, pointing out that any programme seeking greater equality through economic factors will be controversial and contested. Further, disadvantaged groups do not form a single group, being divided internally according to many complex factors. These divisions provide a complex set of viewpoints about possible solutions to the problems of inequality. Similarly, the poor themselves are made up of sub groups, and factors like racism, regionalism and class become important in determining what sub-group aspirations are. Collins emphasises the value of formulating goals, curriculum change, future employment, the work force and political conditions.

The research reported in this section provides some recommendations for pedagogical approaches in bicultural and multicultural classes which mirror the suggestions for effective teaching, and warns against making assumptions about learning styles for different groups in the classroom. The overall message seems to be that good pedagogy is responsive to the individual needs of students, and that in attempting to achieve learning gains for different groups in the classroom teachers need to be aware of and responsive to students' cultures and other factors which result in their being identified as a separate group. An important finding is that teachers should have high expectations for all students.

Summary

The research reports in this section have some consistency with those reviewed in Section 6.1.1. Whilst there is a message that pedagogy needs to be responsive to individual needs this cannot be instead of or at the expense of group needs. Important findings include teachers: having high academic expectations of students; respect for the diversity these students bring; engaging with the knowledge and experiences these students bring to the classroom; developing a pedagogy that is

student centred; and being aware of the teacher-student-community relationship. All encompassing in this research is teacher attitude to the student, their community and their culture for the creation of a safe, inclusive and respectful environment. The literature also warns us against categorising students as non-achievers because of group affiliation, and attaching unquestioningly group characteristics (from anthropological literature) to individual members of the group.

6.1.4 Class size and organisation

A number of studies in the literature surveyed looked at the effects of class size and other aspects of classroom organisation on learning. This research is now briefly reported. It should be noted that these are longstanding issues about which there is a large body of previous literature.

Class size

The central issue is whether class size impacts upon pedagogy, the way curriculum is organised, and student learning outcomes.

Several studies have explored the role of class size in pedagogy. Goldstein and Blatchford (1998) reviewed research into class size effects from a methodological viewpoint, especially concentrating on the strengths and weaknesses of trials and observational studies. The article looked at barriers to learning and points out the difficulties in interpreting much of the considerable amount of research on class size and the effects on cognitive or affective attributes.

Hall and Nuttall (1999) looked at teachers' perceptions of class size and its effects on pedagogy and learning. Class size was seen to interact in complex ways with a range of other variables. The preferred class size of teachers of large classes was considerably lower than the number they were actually teaching. Teachers of small classes judged the ideal class size to be in line with the number they currently taught. The ideal number suggested by teachers of large class sizes was given as the maximum by teachers of small class sizes. This suggests that teachers' existing class size conditions influenced their views about what constitutes the optimum class size. The two groups did not differ in their responses to a series of questions on whether some groups of children and teachers ought to be given preferential treatment regarding class size, and their responses did not vary greatly on the relative importance of class size. Teachers' experiences were concluded to be important when making decisions in this area.

Hattie, in an inaugural lecture at Auckland University (Hattie, 1999) reported a meta-analysis aimed at identifying the major factors that affect student learning. His comment on class size is that appreciable positive benefits were experienced only when classes were cut to 15 or below, with negligible gains for classes over 20 in terms of student achievement, attitude, satisfaction or teacher morale. This effect was the same for primary and secondary schools, across all subjects and ability levels. He suggests that there was little evidence that instruction methods changed when class size was reduced. Most of the improvement could be explained by students getting on better with their work when surrounded by fewer students.

A concern at this summary view is that the link between cultural and other pedagogical factors may not be visible when large numbers of studies are treated in a normative fashion.

Results of other research suggest that there is a link between teacher -student ratio and student achievement. Analysis of the data gathered for the purposes of the 1990-1991 IEA study of reading literacy shows that those countries with smaller class sizes tended to have higher achievement levels of student literacy (Elley, 1994). He notes that this result is masked in the results from countries in which students of lower abilities are purposely aggregated in smaller classes.

Classroom organisation

A number of classroom innovations have been the subject of research found in this literature review. This section briefly reports studies on classroom streaming and integration, teacher collaboration, whole class/small group teaching and resourcing of teaching. The central issue is whether different forms of organisation impact upon student achievement, pedagogy, and the way schools function.

• Streaming and Integration

Classroom organisation is affected by official policies as described in a study by Robertson, Cowell and Olson (1998) of secondary school reform in the province of Ontario, Canada, during the 1990s. This case study, by means of interview and survey, of a secondary school examines the issues surrounding the transition years when two policies, destreaming and integration, were mandated by the Ontario Ministry of Education in 1993, and fully implemented by 1995. Administrators, teachers and students commented on the transition from elementary to secondary school. Particular questions related to destreaming (especially how teachers contend with mixed-ability classes) and integration (with a particular focus on mathematics, science and technology).

The authors found that provincial guidelines for integration were too ambitious and that the teachers felt that they lacked the preparation needed to engage in integration of mathematics and science. Meanwhile mandated destreaming in year 1 of high school occupied most of their time and energy. Students had positive things to say about efforts to integrate but were generally negative about destreaming. The authors suggest that major changes to the culture of the school require a more evolutionary approach rather than that provided by politicised rapid change.

Veenman carried out a meta-analysis of 56 empirical research studies of the effects of multi-grade and multi-age classes. His analyses, (Veenman, 1995; 1996) found that teachers face particular issues when teaching multi grade (level) classes: lack of time for teaching the required content, a greater workload, lack of time for individual attention and remediation, lack of adequate classroom management skills, lack of adequate preparation during teacher training, inadequate materials, and parental concerns about the academic achievement of the children. In spite of these concerns, Veenman found no difference in achievement between multi-grade and single grade classes. The problems of meta-analysis across a very broad field are apparent here. Veenman found that most of the studies he reviewed gave no details about the teaching methods, so it was not possible for him to say whether some methods of teaching multi-level classes work better than others.

Simon (1993) argues, on the basis of case study research, that sorting of students into classes or groups on perceived ability (streaming) has implications for learning outcomes in student access to knowledge, and is also used as a means of social control. This article explores how perceived ability (through TOSCA tests) becomes 'naturalised' because it is perceived to be measuring innate intelligence, and then become interchanged with 'capability' which becomes translated into low expectations of students.

• Teacher collaboration

Friedman examined the relationship between the team concept and school practice using a case study of a team that designed, developed, and implemented an innovative vocational education program within a secondary school. The research report (Friedman, 1997) comments on the effectiveness of teacher collaboration. Friedman argues that the team approach makes sense only if it is accompanied by a shift in thinking about teaching and school practice. This shift involves regarding teams as the primary unit of teaching practice and as a means of linking instructional and structural change within schools. Ultimately a team approach introduces greater uncertainty into teaching and school practice while at the same time providing a means for engaging uncertainty and generating learning.

Friedman argues that unless traditional school structures change, team approaches may not work. Team teaching calls into question the nature of the teaching task itself, so the way teaching is organised cannot be separated from the way the teaching itself is conducted.

Alexander (1995) gathered data about the teaching practices of teachers working in Leeds primary schools. At this time, the Local Education Authority had embarked on an ambitious educational reform process, the Primary Needs Programme (PNP). This programme involved primary teachers in Leeds and was focussed on changing classroom practices and classroom organisation with an emphasis on teaching through group work. The 58 teacher participants were organised into three levels. In level one, 38 were interviewed, and one visit was conducted to see the layout and organisation of the classroom. In level two a further 10 teachers were observed teaching once, interviewed twice, and one classroom visit was conducted to assess the layout and organisation of the room. Level three consisted of an intensive two week programme during which a further 10 teachers were interviewed and observed teaching in their classrooms. Observation of classroom activity showed that teachers organised group work in three ways. The first dubbed by researchers as the “queen bee” approach was one in which the teacher set up the groups, and then placed themselves at a vantage point from which position they directed the group work and answered student queries. The second dubbed the ‘bluebottle’ approach was an arrangement in which the teacher constantly moved from group to group during the lesson, thereby directing each group individually. The third arrangement, which lay somewhere in between these two approaches, was one in which the teacher set groups working on tasks that required little monitoring and worked intensively with one group at a time (Alexander, 1995). Overall, Alexander (1995) found that due to the management difficulties associated with group work, there was a tendency for teachers to set activities that required a low-level of students cognitive processing. Moreover, although the teachers spent nearly all of their time talking with students, individual students spent little time interacting with the teacher.

This disparity between the amount of time teachers spend interacting with students compared with amount of time students spend interacting with teachers was also noted by the ORACLE researchers in 1976 and again in 1996 when the same study was repeated (Galton et al., 1999). However, when comparing data from the 1976 study with those gathered in 1996, the researchers noticed that there was a rise in teacher -whole class interactions. Galton et al. (1999) concluded that “the picture of our typical teacher, then, is of someone who in a class setting alternates between giving information or raising questions whereas when involved with a single pupil he or she is more likely to be acting as a resource (defined in the sense that information the teacher or help from the teacher was essential for the pupil to complete the task) and offering help and feedback on the pupil’s work” (p. 33).

Amongst the results of the afore mentioned project in which 58 primary teachers in Leeds were interviewed and observed teaching, Alexander (1995) found that there did not seem to be a clear correlation between the teacher-student ratio and student time on task. Rather, the time students spend on or off task was related to the individual teacher and the classroom organisational structures under which the students worked. On the other hand, the researchers involved in this project noticed that when there was more than one adult to whom the students could turn for help, there was a marked reduction in the time students spent doing routine tasks such as sharpening their pencils or asking to borrow stationery from other students. This reduction was lost when the number of adults available for assistance grew to more than three at which time the students spent just as much time engaged in routine tasks as they did when only one adult was available for assistance.

In a research report on a project with a complementary focus Pennell and Firestone (1996) discuss teacher networks, a recent innovation in professional development used in California and Vermont to support state policy initiatives. Their analysis relies on interviews, observations and document analysis to identify how teachers' programme experiences were affected by the fit between their backgrounds and programme goals and methods. Drawing on insights from symbolic interactionist

and social movements theory, they treated the teacher networks in these states as professional-area movement organisations sharing the common goal of changing instruction in a constructivist direction. They described the common goals of the programmes in both states, key features of the network programmes and their different policy environments, and how teachers' beliefs, background experiences, social influences, and contextual circumstances affected their programme experiences and willingness to make changes in classroom instruction. The authors argue that those who design professional development programmes must plan for variation among teachers. They suggest that change is not only an apparently rational response to evident problems but a challenge to deeply felt ways of doing things. Networks allowed teachers at different professional stages to discuss how they dealt with the challenges change presents.

• **Whole-class teaching versus small-group teaching**

A meta-analysis of this feature, from a data search that located over 3,000 studies, is documented in Lou, Abrami, Spence, Poulsen, Chambers and d'Apollonia (1996). Their overall finding was that within-class grouping has a positive effect on learning. They conclude that when teachers employ group work it is a useful means of facilitating student learning, especially in mathematics and science classes. Their conclusion is that there are small, but positive effects of classroom grouping; students in small groups achieved more, held more positive attitudes, and reported higher general self-concept than students in non-grouped classes. The optimal group size was small, three to four students. There was, however, some inconsistency across the empirical research studies, so the results have to be treated with caution. An important associated factor influencing sub-class groupings was that teachers were more likely to design different resources and activities for sub-groups, whereas whole class teaching was more likely to employ the same set of materials for all students. Higher student achievement resulted from physically placing groups together and modifying teaching methods and materials.

The composition of groups and their success related to pedagogical approaches. There was a trend that homogeneous ability groups achieved more than heterogeneous groups. However, this needs qualification. High-ability students were unaffected by the type of group, low-ability students did better in heterogeneous groups, and medium-ability students were better off in homogeneous groups. Possible explanations for these effects were that low-ability students received help from more able students in mixed-ability groups. Overall, there was no evidence that one form of grouping was always superior for all students, but there was evidence that under certain conditions different grouping practices lead to different outcomes.

Based on the assumption that cooperative learning is a strategy that produces learning gains for students, Cohen (1994) analysed 94 empirical and theoretical studies in order to investigate the conditions under which small group work is productive. Results of several empirical studies indicate that the teacher should consider the nature of a task before asking students to complete it in groups. If there is not a compelling reason for grouped students to interact, they will complete group tasks individually with little or no inter-group communication. There are three other components of factors that have a profound effect on the productivity of group work: task instructions, student preparation and the role of the teacher (Cohen, 1994). In order to maximise the productivity of small group work the teacher should link task instruction to the expected outcome of the activity. Providing too many instructions may reduce the amount of cognitive processing the students engage in, but too few may mean that students are unable to complete the task. Student preparation should take the form of training for cooperative work so that the group is able to achieve the outcome. Finally, the teacher has to decide on the role they will play during the group work session. Cohen (1994) states that although teachers are often wary of delegating authority during group work sessions, research results shows that student learning is enhanced when they become less directive and more facilitative. Moreover, the degree to which teacher s interact with individual groups also needs to be carefully balanced, because although teacher input can foster interactions amongst group members, it also has the effect of halting the self-directed interactions and communication amongst group members. Thus, for maximum task productivity, factors within the components of task instruction, student preparation and the role of the teacher

should be matched with the nature of the task and the outcome the teacher intends the students to achieve (Cohen, 1994).

• Resources

The relationship between expenditure and student achievement is investigated in a meta-analysis of 60 primary research studies, most from USA, that have investigated this issue. Greenwald, Hedges and Laine (1996) report that it has always been difficult to determine the relationship between expenditure and student achievement even though attempts began over 35 years ago. The authors used two meta-analytical methods: combined significance testing and effect magnitude estimation. Seven input variables were analysed: per pupil expenditure, teacher ability, teacher education, teacher experience, teacher salary, teacher-pupil ratio, and school size. Their general conclusion is that school resources are systematically related to student achievement and that these relationships are large enough to be educationally important. Per-pupil expenditure is strongly related to achievement, as are the three variables of teacher quality (ability, education, experience).

These findings suggest that teachers will achieve better results in schools that are better resourced. Just as importantly, they suggest that teachers themselves are a crucial component of this effect. The better the experience, ability and education of teachers, the better their students' achievement will be, a finding which has important implications for teacher education, both pre-service and in-service. The authors point out that policy makers should ask, how does money matter?, rather than, does money matter? But the issue of resources is very complex indeed, and this study needs to be put alongside other research that time did not allow the project team to process.

Summary

The available literature on pedagogical influences is enormous, and the amount reviewed here is small and, hence, conclusions are difficult to form and are tentative. There is some suggestion that: class size does not necessarily change pedagogical styles of teachers; the research on class organisation regarding streaming and integration is inconclusive; the way teaching is organised is inseparable to the way teaching is carried out; pedagogical practice is more deeply held than is often credited; the research is inconclusive as regards grouping in classes; and resources, including human resources, may affect achievement, but the issue is enormously complex and requires further careful analysis.

Overall summary for 6.1

The literature on pedagogy suggests that this is a difficult area to research because of the complexity of classroom interactions. The research literature is strongest concerning effectiveness factors of teaching, and the importance of teacher knowledge of their subject and of learning. However, there is little research showing direct linkage between these and student achievement. Section 6.1.3 indicates that educational responsiveness to socio-cultural difference is a vast and complicated field, often compounded with the multi-disciplinary and political nature of culturally relevant education. However, the research indicates that it is an area with the greatest potential for change. This research shows that the teachers, their expectations and attitudes, their interactions with the individual student and with their community are central to any changes that need to take place. While organisational aspects of classrooms also feature in the research literature, any conclusions remain tentative.

Overall, pedagogy is an extremely complex area to research, as is the central role of the teacher.

6.2 The role of the teacher

In the accounts of research provided so far in this chapter the role of the teacher has frequently been seen to be crucially important. This intervening and central role in a complex system has been briefly discussed in Chapter 3 and is now considered in more detail in a framework of commentary on the teacher as a key link between the interplay of curriculum, assessment and pedagogy and learning and social outcomes (see the analytical framework in Chapter 3, Section 3.5).

The commentary first considers the role of the teacher in a systemic analysis, then looks further at teacher knowledge and its effects, and finally places research into teacher development in this context. One of the overseas consultants provided a literature review of the central role of the teacher in reform processes when education systems are undergoing change. This section begins with a modified version of this review, followed by an account of further research supporting Olson's review². This commentary provides further discussion of the role of teacher knowledge, and concludes with a survey of research into teacher professional development.

Teachers maintain a balance between their perception of educational needs and the influences from outside such as: government regulations, market forces, educational qualifications, insights from research and cultural norms. Such influences challenge teachers' survival and craft norms (Olson et al., 1998). Challenges to such norms are common themes in the literature across different countries. For example, Pennell and Firestone (1996) found that how teachers responded to mandated state-wide changes in the USA were varied and reflected deeply held beliefs, and Bage et al (1999) found that efforts to achieve system uniformity in the UK in how teachers plan lessons was counter productive due to teacher de-skilling.

Reform can give rise to risks of personal failure, conflict and frustration in a situations where support for self-development is lacking. In the case of imposed curriculum, failure is possible not because teachers lack professionalism, but because of insufficient consideration of teacher qualifications, the existence of different goals for pre- and in-service teacher training, contradictory demands of different stake holders and controversial intended outcomes of a reform. The theme of risk is also common and widespread (Black and Atkin, 1996). For example, Fore (1998) in the US found that imposition of state-wide standards without consultation with professional groups lead to problems in implementation and Spillane (1999) in the UK points to the secure spaces teachers need if they are to be able to radically alter their practice. Anderson (1998) notes that reforms will falter if there are no plans for "collaborative reconstruction of a new social ground".

There are conflicting conceptions of the teachers' role at play in school reform policies, and reform agendas embody images of professional ideals. Teachers usually are outsiders in the policy process. How can teachers participate and enhance their professionalism more actively as subjects and agents simultaneously in this process?

The focus on systemic reform in the title reminds us that teachers and schools are increasingly subject to reformist policies as governments seek to enhance human capital in the face of globalised competition. (See Eisner (1999) for a critique of this globalising trend). Pressures for school reform have intensified as national and multinational test scores become part of the political debates about productivity and the globalisation of educational objectives. Teachers' work is evaluated as never before: measures of teacher capacity are gathered; international studies assess outcomes. These data are used to advance theoretical and political agendas that imply images of teacher professional practice as much they do of student achievement. Often the result of these political appraisals of international results is the call for curriculum reform. Reform is often undertaken and curriculum policy is often used to drive it. How well do these policies work?

Gamoran (1997), in his study of curriculum reform to enhance opportunities for disadvantaged pupils, found that the curriculum policies did affect what pupils experienced in school--opportunity was enhanced somewhat-- but the effects were limited. It was not enough to change the curriculum--teacher practice was also an issue. He argues that unless teacher development parallels reform, the change will be limited. He also points out that new forms of assessment need to parallel curriculum change and that teacher development in relation to both is needed.

² A review of recent literature on change in the context of moves towards systemic reform. Prepared by John Olson, Queens University, August 1999.

These kinds of conclusions arise across many nations. For example, in the recent OECD report (Black and Atkin, 1996; Raizen and Britton, 1997) - it was found that there were diverse approaches to change ranging from support for teacher development to systems of imposed policies directed by educational objectives. However, whatever the reform intentions from outside schools were in the different OECD countries, they are modified by teachers as they enact those intentions; often quite radically; and other studies have found the same (Tyack and Cuban, 1995).

There are many pressures on teachers to reform practice. The teaching profession is increasingly challenged by accountability measures adopted by governments. Systemic reform can be seen as a response to these pressures as well as a reflection of what has been learned about the reform process over the last three decades. Measurement of outcomes is a significant part of the systemic reform process; gathering knowledge of the effects of changes in the system (Greenwald, et al. 1996). Establishing specific outcomes and mandating them is part of the way change is pressed upon schools. Results from measurement of outcomes are used to shape the system and to put pressure on it. There are dangers here. Haertel (1999), for example, notes that measurement-driven reform strategy is problematic because high-stakes assessment is founded on a number of untested assumptions. For example, he notes that portfolios used to assess performance have limited generalisability. If the measurement instruments are weak then the system lacks credibility.

Strident reform ideologies, disappointing scores on international tests for some and a climate of distrust of professionals generally, to mention only a few trends, challenge teachers to account for themselves by taking up the basis of reforms and questioning their validity in a dialogical way. How are teachers able to assess and question the system? Attention increasingly is drawn to concepts of self-development, self-assessment, deliberation, reflective practice and collaborative partnership to support teacher participation in practice as well as policy processes. These professional practices entail, questions about human values, beliefs and moral considerations. Teachers need to be able to bring such issues to the policy table in the context of systemic reform.

Unfortunately reforms usually are not based on the insights teachers have achieved from practice over time, as the recent OECD study has underscored (Black and Atkin, 1996). Atkin (1998) notes that much can be learned from the OECD study about systemic reform. He says that it was evident that teachers played a central role in the 23 reforms studied in the OECD project. Teachers are able to provide input to reformers about the dilemmas of reform. While governments might be concerned about learning outcomes and their political consequences and thus call for reform, more needs to be known about what happens if teachers do not understand and support reform.

Teachers do not work in isolation. Those who work with them are caught up in the reform process. The professional self is developed in a community of persons involved in teaching. The process of collaboration in innovation is part of practice and its reform. What does research say about collaboration?

Hatch (1998) studied how teachers collaborated in a major curriculum reform project in the USA. He found that differences in teacher theories of action existed in the group and that these differences hindered collaboration. Unless these differences are surfaced and discussed these hidden theories will exert major but poorly understood influences on the collaborative process. Friedman (1997) studied the effectiveness of teacher collaboration. He argues that unless traditional school structures change team approaches may not work. Team teaching, he notes, calls into question the very nature of teaching itself.

Robertson et al (1998) found that teachers working on an integrated math/ science high school curriculum were challenged by the task of teaching across these subjects and tended to retreat to their subject domain where they felt they had control over the pedagogy. A common ground for looking at these subjects in concert was not found, partly because these teachers were coping with destreaming (detracking) at the same time. Too much reform pressure was at work to enable them to find the space to consider how their subjects connected. While they may have been able to make

progress, given reflection on their experiences, there was no provision for this to happen. Without the needed supports for reflection--that is to say without some kind of systemic effort--reflection remains a wishful nostrum.

The value of current models of teacher reflection as a professional development strategy is being questioned. Waks (1999) notes that the work of teachers is not well represented by using comparisons with design-based professions such as urban planning. Reflective practice-- such as those adopted by these professions-- are not comparable with what teachers do as they learn from experience, he argues. Teachers learn from experiences in actual classrooms. There is furthermore the more general tendency--reflected in models of reflective practice and teacher knowledge alike-- to equate teacher expertise with cognitive capacity while undervaluing the collective dimensions of the practice which are sustained through the virtues teachers bring to the profession (Olson, 1992; Hansen, 1999; Jackson, 1999).

The emphasis in reflective models on thinking in action and after it may discount the importance of the thought that goes into the planning beforehand. Carlgren (1999), for example, notes that there has been an overemphasis on thinking about teaching after the fact. What about the planning that goes into teaching? What is the quality of that? How well is that done? How much does that planning reflect what teachers have found out from prior experience? Here Carlgren is asserting the importance of didactics; an aspect of teachers' work often left out in models of teacher reflection.

Furthermore is the reflection even if undertaken effective? There is more to reflection than cognition. Klette (1997), points to the virtues needed to accept what one learns from reflection. She is sceptical about how well teachers learn from reflection. When confronted with evidence of a lack of success, teachers may be inclined to romanticise their work--to discount its realities--or simply ignore them. The difficulties of reflection raise questions about the nature of in-service education that might be needed to confront the challenges of systemic reform. These difficulties point to the virtues needed to learn from one's experience.

What role do consultants and supervisors have to play given the difficulty of reflection and the demands of reform? Greater collaboration amongst these groups and researchers seem called for with less an emphasis on perfect implementation than on learning more about practice and the possibilities of reform from some dialogue between them. It can be argued that if these groups could develop common didactical ground then systemic reform might be more productive. (For a model of the pedagogical and didactical work of the teacher in the context of reform see Olson, et al (1998)).

One clear theme in reform is that teachers are asked increasingly to offer students more individualised and individualistic instruction. The point of view of the students is increasingly sought in research and in recommendations for practice. In this context of greater attention to student interests and differences, teachers are increasingly expected to act as clinicians and managers of new technologies as they implement rapidly changing educational policies. Teachers may find that they cannot accommodate these increased demands. For example, Griffin (1998) in Australia found that teachers were expected to give detailed reports about students. They found this onerous and without much benefit on the face of it, at least in their view.

Changes in policy, research questions and school practice, as we saw, challenge existing definitions of professionalism. The question of professionalism has always been a thorny one for teaching: when compared to medicine and law teachers have often been assigned to the status of a semi-profession. The degree of independence and specialised knowledge deemed necessary to qualify as a profession is said to be absent in teaching. Yet teachers are asked to take charge of the development of the young from an early age and over many years. It is an onerous responsibility. Teachers do this work largely unsupervised and mostly according to their own lights, efforts to manage their work through curriculum and supervision notwithstanding.

Teachers are called upon to be professionals yet there is much in the process of systemic reform that may call into question this professionalism. The role of supervision of teachers arises in the context of systemic reform as elements of the system are assessed for success in implementing system wide changes. The danger of an over zealous use of supervision--or an overly standardised approach--is that teachers respond differently to different forms of supervision and here again the urge to create system-wide standards may over-ride approaches to supervision which are sensitive to individual differences . For example, Zepeda and Pontecelli (1998) found that supervisors need to know how teachers make use of critical feed back. The system of giving the feed back may be uniform but how it is used is not. This question of individuality arises in relation to collaboration, as we saw above, and it arises in the way teacher respond to challenges to their preferred forms of pedagogy.

There is a danger, however, of over-emphasising individual teacher differences. Teachers share traditions of practice and these play an important role in what happens in reform. These traditions are part of the collective wisdom and need to be understood. Efforts to get around teacher tradition have not been successful(Tyack and Cuban, 1995). Take, for example, efforts to drive change through student assessment . Such assessment-driven change comes up against the beliefs that teachers have about what they are teaching and how they teach it. These beliefs are varied and deeply held as we saw. Cheng (1999), for example, found that an effort to change how English is taught by changing the nature of the external examinations ran up against core beliefs of teachers. While teachers modified their practices to accommodate to the exam, few of their core beliefs changed.

While systemic approaches understand that the interaction amongst factors influencing change are complex and that piece-meal change is problematic, there is a danger that the press for uniformity will be counter productive. First, the changes may not speak to issues that concern teachers differently in different parts of the system. School cultures are varied within systems. Secondly, while central administrators may have theories well grounded in social science to work with, there are traditional theories that teachers hold which are well adapted to the conditions of practice but which are not well codified. Thirdly, teachers need to be able to recover these theories in the context of systemic change and find ways of using them to engage the changes they are faced with. Finally, more dialogue amongst professional groups is needed if systemic reform is to work. These points comprise much of the research mentioned in this review.

Research coherent with Olson's paper

In research commentaries which recognise the complex systemic nature of change processes (Darling-Hammond, 1994 (a) and (b); and Darling-Hammond and Falk, 1997) the argument is advanced that instead of starting with content and performance standards, policies should aim to create a system in which improved teacher knowledge and equalised school capacity are the starting points for systemic change. In such a system, teachers and schools will have the knowledge, resources, and organisational supports to create appropriate curriculum and useful assessments for the students they serve.

Fore (1998) studies the implementation of new curriculum standards in the state of Virginia. There was considerable debate about the standards. They were imposed without consultation with professional groups. This study points to the dangers and drawbacks of systemic reform without participation of key stakeholders.

Friedman studied the effectiveness of teacher collaboration (Freedman, 1997). He argues that unless traditional school structures change, team approaches may not work. He also points out that team teaching calls into question the nature of the teaching task itself, so the way teaching is organised cannot be separated from the way the teaching itself is conducted.

Lee and Smith (1996) investigated how organisation of high school teachers' work lives affects their students. The authors concluded that schools where most teachers take responsibility for learning are more effective and more equitable environments. The researchers noted that it is difficult to change teachers' beliefs about student learning.

The research into authentic pedagogy in Dutch secondary education (Roelofs and Terwel, 1997) has already been discussed. In the context of Olson's paper the lack of support provided for teachers is relevant.

A study which revealed the complexity of interactions between curriculum, assessment and pedagogy when change is introduced is reported in Lawrence and Tatum (1997). The authors report that an anti-racist professional development programme for white teachers affected their curriculum content and delivery, their relationships with students and community, and motivated institutional reviews of student assessments.

Teacher knowledge of the realities of the classroom

Olson's review discusses the professionalism of teachers and their increasing role in managing rapid change. He notes that decisions about change are not always made with referral to this accumulated knowledge of the teaching profession, and that the changes could benefit from such a process. This comment is mirrored in other research reports.

Atkin (1998) reflects on what can be learned from the OECD study about systemic reform. He emphasises the strong role teachers played in the reform process in different OECD countries, and the importance of supportive conditions for teachers. System change, he concludes, can occur but it needs to involve the wisdom of those who have experience of the dilemmas of changing well-established practices.

Galton, Hargreaves and Comber (1998) discussed why teachers change or do not change in the face of new curriculum development. They comment that central authorities making curriculum decisions seem to view curriculum change and associated changes in classroom practice as a 'rational' activity; when change is prescribed it will happen in schools. The authors comment that this is far from reality. Changes in curriculum and teaching strategies need to involve teachers in their development.

In a further analysis (Galton et al., 1999) of the effect of the National Curriculum in England and Wales on primary teachers, a shift in the culture of primary teaching is described. The authors argue that teachers used to see teaching as a way of life rather than just a job and were prepared to work long hours to this end. Their professional judgment was felt to be respected. They saw teaching as made up of the application of principles and skills in a flexible manner, taking into account the particular group of students. Imposed changes have caused difficulties, for example when teachers were required to teach 60% of whole class teaching during mathematics lessons. Teachers were resistant to some of the imposed changes because they saw them as threatening to their understanding of their professional responsibilities.

In a major review of primary teaching and learning in England and Wales already referred to in Section 6.1 (Gipps, 1999) the issue of the complexity of pedagogy and enforced changes to classrooms is also explored. Gipps comments that we need to talk in terms of teachers having not a pedagogy but a range of pedagogical strategies (which includes a range of materials and content, teaching style, and classroom rules/agreements) to be used at different times. One quarter of the teachers in her research study felt pushed into doing more transmission teaching because of the amount of work to be covered by the National Curriculum, or in some cases by tests. They felt unable to use children's ideas as starting points for work and regretted this. Some teachers felt they had lost control of the curriculum, and others said they stuck to their own beliefs and practices in assessing children.

The professionalism of primary teachers in the UK has been engaged in their informal assessments which contribute to national assessment. Gipps, Clarke and McCallum (1998) studied informal assessment processes used by over 300 teachers. While teachers found their part of the assessment programme very time consuming and demanding, they considered teacher input essential because of its direct impact on student learning. The study provides evidence to suggest that teachers' skills and knowledge of assessment procedures have improved since teachers have been part of the process. Teacher participation allows teachers to have a stake in the process, and ensures that a broad range of skills and processes are assessed. In an argument similar to that in Galton et al (1998; 1999) the authors point to a need for professional development and resources to support the changes.

Gitlin and Margonis (1995) make a further contribution to this issue when describing the need for teachers to be taken into the confidence of curriculum designers. Curriculum and other reformers are described as having underestimated the difficulties caused by fundamental changes in teachers' work. They describe teacher resistance to change as not necessarily negative, since it may be their professional judgment that the likely impact on students and student achievement will be negative, and they may then appropriately modify the intentions of the change.

Hall and Nuttall (1999) refer to the need for teacher personal knowledge of how class size interacts in complex ways with a range of other variables to inform decisions being made about class size.

Based on interviews with teachers working in networks in two American states, Pennell and Firestone (1996) report that teacher background, social influences and contextual circumstances coloured their experiences of networking and their willingness to make changes. They argue that those who design professional development programmes must plan for variation among teachers. They suggest that change is not only an apparent rational response to evident problems but a challenge to deeply felt ways of doing things.

Summary

Olson's commentary argues that teachers do not have enough input into major reforms and that this suggests that the professionalism of teachers is undervalued. He stresses that reform will not happen unless teachers support the reform and are involved in the decision-making and the change processes. This requires a dialogue between teachers and other participants in the reform process. Furthermore, he argues that if reform is to take place it has to be done alongside teacher development, as reform often involves risks and major changes in pedagogy which can be resisted if change is not well managed. The research literature supports Olson's analysis.

6.3 Teacher professional development

The teacher is clearly central to improving student achievement. The continuing education of the teacher is linked to curriculum change that may influence achievement. A significant focus of John Olson's essay (above) is that teachers as professionals are both subjects and agents of change. He notes that they are too often outsiders in curriculum development, in the development of assessment processes, and in the pedagogical approaches implied or required in these changes. Teachers' work is evaluated as never before: measures of teacher capacity are gathered; national and international studies assess outcomes. These data are used to advance theoretical and political agendas that imply images of teacher professional practice as much as they do of student achievement.

Olson argues that teacher development in relation to both curriculum change and changes in assessment regimes is needed (citing work by, for example Gamoran, 1997; Black and Atkin, 1996; Raizen and Britton, 1997). He sees as a major question: How are teachers able to assess and question the system? Teachers need to be able to bring their knowledge to the policy table in the context of systemic reform, and to be supported through professional development processes when major changes occur.

Unfortunately reforms are not usually based on the insights teachers have achieved from practice over time, as the recent OECD study has underscored (Black and Atkin, 1996). Atkin (1998) notes that while governments might be concerned about learning outcomes and their political consequences and thus call for reform, little will be accomplished if teachers do not understand and support these reforms.

Olson cites many examples of the need for teacher professional development. Friedman (1997) studied the effectiveness of teacher collaboration. He argues that this change from the traditional school structure needs support through development programmes.

Griffin (1998) in Australia found that teachers were expected to give detailed reports about students. In the absence of assistance through professional development they found this onerous and without much apparent value.

Galton, Hargreaves and Comber (1998) discussed how the introduction of the national curriculum in England and Wales affected assessment processes used by primary teachers in small rural primary schools. They concluded that the new requirements for whole class teaching need associated professional development support for teachers.

Sizmur and Sainsbury (1997) point to the need for in-service training when developing national assessment procedures.

The results from extensive professional development can be slow in coming. Sheppard (1995) reported 2 years of professional development research in Colorado where teachers met after school every week in workshops run by the researchers. The project required a large amount of time to develop teaching strategies for improved learning in mathematics and reading. They report gains in student learning in mathematics but no change in reading. They also note that professional development needs to be extensive and in depth, a requirement that is often underestimated.

Bell and Cowie (1997) report gains in teacher understanding of pedagogical approaches with ten teachers over two years when they introduced formative assessment into their classrooms. The teachers reported their involvement in the research as a professional development opportunity.

Complex change can cause problems. Robertson, Cowell and Olson (1998) report a case study of a secondary school in Ontario, Canada where two policies, destreaming and integration, mandated by the Ontario Ministry of Education in 1993, and fully implemented by 1995, were examined by means of interview and survey. They found that the provincial guidelines for integration were too ambitious. The teachers felt they lacked the preparation needed to engage in integration of mathematics and science, when the simultaneous mandated destreaming occupied most of their time and energy. Students had positive things to say about efforts to integrate but were generally negative about destreaming. The authors suggest that major changes to the culture of the school require a more evolutionary approach rather than that provided by politicised rapid change. This process may well have been more effective with professional development support.

Pennell and Firestone (1996) argue that the design of professional development programmes must plan for variation among teachers. They suggest that change can challenge deeply-felt ways of doing things. Networks allow teachers at different professional stages to discuss how they deal with the challenges change presents.

The change process can be mandated as described by Harris and Drake (1997), who report on a high school that initiated reform using school wide action research teams as a vehicle to develop teacher leaders, teachers as change agents, and a collaborative reflective culture. The concept was introduced by the principal who mandated that all must participate. In the first two years teachers worked together to make meaning of the action research process. As they gained experience, they became more confident. In the third year, the teams finished individual projects. There were both

positive and negative aspects of this programme. While most teachers felt they had benefited from their work on the teams, they also reported that the time-consuming nature of this work impacted negatively. One issue which came up repeatedly was that mandated collaboration was problematic since there were not shared goals for the type of change and development. Without shared goals, some teachers did not do their share of the work. A lack of commitment by some teachers caused some obstructive behaviour and resistance. The principal justified the mandatory participation by saying that there was no other way to begin on the challenges for change faced by the school.

Bell (1993) describes a three year research project funded by the New Zealand Ministry of Education which reported a number of factors affecting teacher professional development. The project team investigated the learning of some teachers of science in four teacher development programmes. Programmes were run to help teachers change their classroom activities to take into account international findings of the research about students' learning, and in particular the findings of the previous Learning in Science Projects. To do this, the teachers had to take into account students' thinking as part of their teaching. The findings indicated that teachers changed their concepts about learning, teaching, the nature of science and the curriculum; their practice with respect to their activities and roles in the classroom; and their feelings about aspects of science education. The teachers also developed their ideas of what teacher development is and what promotes it. The main finding was that teacher development involves professional, personal and social development. Factors that influenced the teachers' development were talking with other teachers and using anecdotes; a focus on teaching in the programmes; teachers feeling better about themselves as teachers; better learning conditions and outcomes; teachers being able to contribute to the programmes; support; feedback; reflection; managing school restraints; teachers' prior ideas, experiences, concerns and feelings; components in the programmes such as journals, new teaching activities to try out and sharing sessions; facilitation; and teachers becoming empowered to take responsibility for their own development. A model of teacher development was developed.

Pearson and Bell (1993) indicated that the teachers in their teacher development research project developed their teaching in two ways. Firstly, these teachers used different activities in the classroom such as concept mapping, brainstorming, a post-box technique, eliciting students' questions as an interactive teaching approach, small group discussions, interviewing, open-ended investigations, and card-sorting. These activities were suggested in the programme so that teachers could take into account student thinking. The teachers often had to change their practice and thinking about student behaviour management, resource management, assessment, covering the syllabus, and recording and reporting the learning outcomes. This reconception of their role helped teachers to move from teacher-centred and teacher-dominated activities to more student-centred activities. They were still in control but in different ways.

Secondly they were able to find out what concepts, opinions, concerns and interests students brought to a lesson; what the students were thinking during the lesson; what understandings the students were constructing during a lesson; and what learning the students had achieved. Most importantly, they were able to develop the ways they interacted with their students' thinking.

Carlgren (1999) addresses the issue of teachers' professional knowledge and professionalism in relation to the growing design aspects of teachers' work. The teaching profession is described as a profession characterised by ruptures as a consequence of school reforms. New demands on teachers are made as a result of the reforms and, as a consequence, the character of the teaching profession is changed. In a comment on reflective practice, her view is that teacher knowledge which has been embedded in practice must be replaced by a 'disembedding' of this knowledge. She says that what teachers do before they teach is just as important as reflection during teaching.

In a similar vein Walsley (1999) points towards professional development as a key to increasing teachers' repertoires, recommending professional development alternatives to the usual practice of one-off sessions with no follow-up, techniques imported from outside the school, only one participant in programmes from a school, and the assumption that every teacher would master the techniques. The paper recommends that teachers need a confirming, complex image of teaching; an

interconnected system of support; adequate support for emerging teachers; mentoring; on-going teacher development; and teacher leaders. They describe teacher professional development as ‘a place to advance’ in teaching.

Lawrence and Tatum (1997) report that an anti-racist professional development programme for white teachers affected their curriculum content and delivery, their relationships with students and community, and motivated institutional reviews of student assessments.

Gipps, Clarke and McCallum (1998) report that teachers need much more help in assessment processes which can best be provided in the form of exemplification materials and some form of group moderation.

Summary

The literature documented in this section provides research evidence that professional development is an integral part of any major reform including curriculum and/or assessment. There is strong evidence that professional development associated with major reforms needs to be extensive and in depth, particularly since aspects of teaching may be held deeply and personally and reform may challenge these beliefs. While some research shows that professional development may make some gains for teacher practice associated with the reforms, change to strongly-held beliefs is likely to be slow and require support after the development programme. Longer programmes can lead to teacher frustration at the time required, so other issues arise. When teachers consider changes to have been imposed on them they often resist the reform. Professional development processes require further research to clarify these issues.

6.4 Effects of pedagogy on students’ learning and social outcomes.

Many comments on the interplay between pedagogy and learning outcomes have been made in the research described above. The influences of curriculum and assessment on these learning outcomes has been represented in the analytical framework for this review (Figure 2), and the difficulty of finding direct links referred to at the end of each of the previous two chapters. The next chapter brings together the research on student learning and social outcomes using the analytical framework for this review.

General summary

It would appear from the literature that the role of the teacher is central to any research on pedagogy as a result of the intervening role of the teacher in the classroom. The research literature about pedagogical practice provides information about:

- Effectiveness factors of teaching include the following characteristics:
 1. having a broad understanding of curriculum aims and objectives;
 2. having a wide range of pedagogical strategies;
 3. having high expectations of all students;
 4. knowing their students well;
 5. providing good feedback to students;
 6. recognising student success;
 7. having sound content knowledge of their subject; and
 8. taking responsibility for the progress of their students.

- Findings on successful teachers of culturally different groups centre on attitudes. These teachers were found to:
 1. have high academic expectations of students;
 2. respect the diversity these students bring to classrooms;
 3. engage with the knowledge and experiences these students bring to the classroom;
 4. use a pedagogy that is student centred; and
 5. be aware of their role in teacher-student-community relationships.

- It is difficult to quantify teacher effectiveness, since decision making about pedagogical strategies in classrooms is a complex professional response to many factors.
- Teacher subject knowledge and pedagogical knowledge can affect planning, assessment, implementation of curriculum and curriculum development. It appears that low knowledge levels may restrict classroom practices and learning opportunities for students.
- Conclusions about class size and organisational factors are tentative:
 1. class size may not necessarily change pedagogical styles of teachers;
 2. the research on class organisation regarding streaming and integration is inconclusive;
 3. the way teaching is organised is inseparable from the way teaching is carried out; pedagogical practice is more deeply held than is often credited;
 4. the research is inconclusive as regards grouping in classes; and
 5. resources, including human resources, appear to affect achievement, although quantifying the affect is difficult.
- Research literature supports Olson's analysis that any major reform needs teacher professional input, and teacher development needs to parallel the reform.
- There is a strong argument that professional development to support reform needs to be extensive and in depth, but reformers need to take into account a number of complex factors.
- Teachers can hold pedagogical practice very deeply and personally, and professional development has not always addressed this significantly.

Chapter 7: Synthesis of Chapters 2, 3 and 4

Introduction

The previous three chapters have looked at curriculum, assessment and pedagogy, seeking to describe the effects of these factors of schooling on learning and social outcomes. The analytical framework for the review (Figure 2) gave an overview of the interactions between the three nodes of curriculum, assessment and pedagogy and represented how learning and social outcomes were mediated by the roles of the teacher and the learner, most significantly where they come together in the classroom. The analysis so far has supported this focus for this review.

The complexity of the interactions means that students' learning and social outcomes are the result of many factors. The first focus in this analysis was the curriculum.

Curricula are decided through the interactions of many interested parties (or pressure groups) in the society for which they are designed. Their definition can assist teachers by providing a framework of the material to be taught. Such a framework assists formative assessment by indicating the pathways for learning, though the progression in curricula are not firmly based on research, Black et al (1996). A powerful contribution to learning can come from teachers working together to design programmes for good learning in the classroom. One effect of the introduction of the new national curriculum in England and Wales was that teachers responded by developing processes together, and by professionally trying to implement this framework of content and skills. There is evidence that the initial reaction of many teachers was to resist aspects of the curriculum which were not coherent with their views of good teaching and learning (though this effect was considerably influenced by the assessment procedures which were introduced concurrently with the curriculum). This response stemmed in part from the teachers being insufficiently consulted in the curriculum development process. The account of this innovation continues with less positive effects. Substantial changes were made, and frequently, to the new curriculum and accompanying assessment processes, and teachers became more resistant and more professionally dissatisfied with a process and product which continued to lack consultation with them.

The framework provided by a curriculum can then have beneficial effects on the classroom, but development of this framework needs to be in consultation with the teachers who will implement it in their classrooms. Key factors in this consultation would be questions about the definition of content and progression (this review does not find research answers to the question of what should be in the curriculum, and how detailed the description of this content should be), and how teachers' professional knowledge of students' learning processes can be integrated with the curriculum.

The second factor in the analysis was assessment.

Assessment has been described in this review as a major influence on curriculum implementation and consequently on learning. The intentions of a curriculum can be distorted by an assessment process which gives high stakes to some aspects of the curriculum, or which narrows its potential. Summative and formative assessment processes each have their place, but it is vital to understand the contribution each can make, and the problems which arise if summative assessment inappropriately dominates. The research evidence which describes the benefits of formative assessment processes for learning in the classroom provides a major indication of potential benefits to classroom teaching and learning, provided that this form of assessment is not swamped by high stakes summative assessment.

The third factor in the analysis was pedagogy.

Pedagogy has been subjected to considerable analysis and research. The review describes pedagogy as strongly interactive with curriculum and assessment, and notes that important consequences for learning flow from the way that curriculum supports or hampers teachers, and the

way that assessment processes assist the teaching and learning process or stand in the way of good practice. Continuing professional development is seen as vital to support teachers in their implementation of innovations, and to inform their contribution to these developments. Their knowledge of the subject they are teaching, and of their students; and their expectations of and interactions with students are crucial to learning and social outcomes.

The complex interactions between curriculum, assessment and pedagogy define a system which requires approaches to its understanding which recognise these interdependencies. The classroom is where all these nodes come together and is the centre of the review focus. What happens in the classroom is dependent on the professional role of the teacher in responding to and interpreting the system. The classroom is also where the learners' many abilities, attitudes to learning, and cultural background have their effect. Learning and social outcomes are greatly affected by the systemic interaction and the intervening role of the teacher is crucial (it was noted in Chapter 3, Section 3.4 that influences outside the classroom make another important contribution).

The preceding chapters have, then, provided an elaborated commentary on the analytical framework outlined in Chapter 3. In the following section the research about learning and social outcomes is discussed.

7.1 Learning and social outcomes

A number of analyses of the factors influencing student achievement have been attempted. This commentary first looks at general overviews and then shifts the perspective to detailed studies.

Hattie (1999) in an article based on his inaugural lecture at Auckland University, summarises over 400,000 studies covering 50 million students to identify the major factors that affect student learning. The highest effect is described as reinforcement, and other high ranking effects are: students' prior cognitive ability, instructional quality, instructional quantity, direct instruction, acceleration, home factors, remediation/feedback, students' disposition to learn, class environment, challenge of goals, bilingual programmes, peer tutoring, mastery learning, teacher in-service education, parent involvement, homework, questioning, peers and advanced organisers. Testing, learning hierarchies and finances are ranked, along with other factors much lower in their effects.

There is considerable coherence between this list and the overview to this chapter above. Hattie recommends "dollops of feedback" providing information on how and why the student understands and misunderstands what directions the student must take to improve as a powerful tool to enhance achievement. He also sees as important the teacher setting challenging goals and structuring situations so students can reach these goals.

Hattie's comments fit with the analytical framework for this review in recognising the many influences on learning, with the role of the teacher being central.

The complex issue of what constitutes the most appropriate pedagogy is explored with a different focus by Phillips (1997) who argues that high expectations for student achievement, clear achievement oriented goals, the amount of time spent on instruction, and the amount of homework assigned to students (described as academic press), have a strong relationship with academic achievement in mathematics. This seems to fit with Hattie's comments. However Phillips's analysis goes on to contrast academic press with approaches which first aim to satisfy students' and teachers' social and personal needs before learning can take place effectively (described as communitarian). The author found little support for the hypothesis that communitarian climate enhances attendance or mathematics achievement. He found that in schools where the average level of teachers caring for students is relatively high, students' test scores were relatively low. These findings suggested to him that teachers in some schools may be more concerned with maintaining effective relations with students than with imparting skills and knowledge, and that teachers who focus on students' social and emotional needs may neglect their academic needs.

Evidence for his argument comes from students being less likely to be absent when they attend schools where a larger proportion of 8th graders take algebra, and where teachers expect them to graduate for high school and college. Similarly, students were found to learn more about mathematics during their middle school years when they attended schools where students do more homework. This commentary points to the complexity of measuring the interaction between teaching approaches and learning.

Gipps and MacGilchrist (1999) in a literature review of determinants of good learning, make the following comments (which are in some disagreement with Phillips and more in agreement with Hattie):

- Socio-economic inequality is a powerful determinant of differences in cognitive and educational attainment. Social class, along with ethnic background, gender and disability, has been found to have a substantial influence on the life chances of young people.
- Schools can and do make a difference, but some schools are much more effective than others at counteracting the potentially damaging effects of disadvantage.
- Teacher expectations influence pupils learning. Characteristics of highly successful schools include the quality and nature of the teaching and learning in classrooms along with the overall learning ethos of the school. In particular, the relationship between teachers beliefs and attitudes and pupils progress and achievement.
- The literacy and numeracy knowledge and skills that children had acquired before they started school were a strong predictor of attainment at age seven. But school factors that appeared to exert a greater influence on progress than home background were the range of literacy and numeracy taught to the children and the teachers' expectations. The school-based measures of teachers' expectations of children's curriculum coverage showed the strongest and most consistent association with school progress. Where teachers had low expectations of children, they provided a narrower curriculum offering.
- Low teacher expectations produce low motivation in students to learn.
- Optimal conditions for learning include a positive, personal learning attitude where challenge is high and anxiety and self-doubt is low. Recommendations are for plenty of praise and encouragement, high levels of expectations of all children and high levels of work-related talk and discussion.

In summary, teachers who see learners as thinkers, and capable of achieving more, enable children to view themselves as able to learn.

Reports which broadly agree with aspects of Gipps analysis come from Lee and Smith (1996), Ross (1995), Slack (1998) and Slavin (1995).

In a major study of secondary schools Lee and Smith (1996) investigated how organisation of high school teachers' work lives affects their students. The authors conclude that schools where most teachers take responsibility for learning are environments that are both more effective and more equitable. Because the results take into account differences in schools social and academic characteristics, as well as the academic and social characteristics of their students, the authors believe that this represents a solid finding about the social organisation of schools. The authors also considered the cooperative nature of teachers' professional community, and concluded that schools with high levels of cooperative and supportive activity among the staff are places where students learn more in some subject areas, and where learning is somewhat less stratified by social class. Cooperation among teachers made schools both more effective and more equitable environments.

Ross (1995) argues that teacher efficacy arises from goal-setting and attributional processes. Teachers who anticipate that they will be successful set more challenging goals for themselves and their students, accept responsibility for the outcomes of instruction, and persist through obstacles. The author argues, from a research review, that changing teacher beliefs about their effectiveness will have a positive effect on student cognitive and affective achievements because they are more likely to set higher goals to attain.

Slack (1998) sought to determine whether students in accelerated schools with distinct contextual features attained significantly different test performances from students in regular schools. Results were determined using the Louisiana Educational Assessment Programme (LEAP) and the California Achievement Test (CAT). The study suggests that disadvantaged students have most to gain from innovative teaching approaches. Over 70% of lower ability students improved test performance. Higher ability students were less likely to improve. Changes to test scores were attributed to longer periods on manipulative material and the general climate or ethos of the accelerated schools.

Slavin (1995) surveyed the research literature on co-operative learning strategies, in groups that reflected class constitution in the US. In 16 of the 19 studies reviewed positive effects on intergroup relationships were found. The research also indicated positive effects on student achievement in a wide variety of subjects and for students of different ethnicities and backgrounds. The results on how much positive effect and for whom were mixed. Some studies showed an equal gain for white students as well as African American students, while others showed unequal gains between the groups.

A student view coherent with the Gipps analysis is provided by Morris (1995).

This study explored the in-school and out-of-school factors which the learning of senior boys and girls in the secondary school setting. The researcher administered questionnaires to over 400 students and reported students' views about the effectiveness of their teachers. The best teachers were seen by the students to be skilled communicators who knew their subject and were personally interested in each student. They were not remote transmitters of knowledge but people who could establish rapport with adolescents and still maintain their professional distance.

A report which both complements and disagrees with Gipps' research comes from Levine and Lezotte (1995). The authors develop a set of characteristics for "effective schools" which include: productive school climate and culture reflecting shared values; safe and orderly environment; teacher commitment to a shared and articulated mission focused on improving achievement; problem solving orientation; staff cohesion, collaboration, communication and collegiality; staff input in decision-making; school-wide emphasis on recognising positive performance. Teaching staff are seen to be pivotal to successful schools; the criteria fit well with Gipps (1999) with the emphasis on active and enriched learning in an interactive context. However, the authors also report that effective pedagogy for low achieving students comes from a very structured environment (such as specific grouping arrangements) and focused delivery (along the lines of mastery learning).

In similar vein to Levine and Lezotte's analysis about specific grouping for low achieving students, Gamoran and Weinstein (1998) investigated 24 highly restructured schools in the US. The central issue was whether providing students with common experiences, or addressing differences among individual students affects overall achievement. Whereas many schools typically divide students for teaching into different tracks by ability groups, the tendency in restructured schools was to minimise differentiation (called detracking in the article). The authors report that moving away from ability classes in this way seemed to be more difficult in secondary schools than in elementary schools. In mathematics, ability classes still seemed to prevail. The study found that in some schools detracking had a negative effect because it seemed to lower performance of all pupils.

Hattie, Biggs, and Purdie (1996) aimed to identify features of study skills interventions likely to lead to learning success. A meta-analysis of 51 studies looked at interventions to enhance student learning through one or a combination of learning or study skills - such as self-management of learning, or effective components such as motivation and self-concept. The results support the notion of situated cognition, recommending that training other than for simple mnemonic

performance should be in context, use tasks within the same domain as the target content, and promote a high degree of learner activity and metacognitive awareness.

The meta-analysis suggests a clear and comprehensible picture emerging. If the intention is to teach for simple retention of accurate detail, then the use of mnemonics - such as using imagery or linking items to be learnt - is highly effective. If the intention is to help students understand content with a view to applying it in a new context, then more complex strategies are indicated.

The thrust of the findings is quite compatible with the thrust of situated cognition and systems theory. That is, improving learning is less likely to be achieved by targeting the individual in terms of a deficit model, which presupposes that the individual is lacking the right strategies and needs to be taught them or is using the wrong strategies and needs to have them removed.

On the issue of the use of computers in the classroom, Hattie (1999) reports a very large review of 31 meta-analyses, 17952 studies, and 352 effect-size studies to demonstrate that computers in classrooms did make a difference. Compared to classrooms without computers, students in classrooms with computers improved their achievement by approximately three months or 15%. Hattie argues that these gains are actually quite small, as is the variability within and between groups; for example, the effects decrease with age, with primary students gaining most, but secondary students gaining less. There were effect changes on gender between secondary students but not primary students. An important outcome, according to Hattie, is that an innovation like computers in classrooms should be compared to other innovations to see which give the best improvement. He says that the evidence from his analyses are that most innovations result in similar gains to those in the computer example, that is, small gains, with most of them smaller than the computer introduction. Some examples are: programmed learning, use of calculators, testing, ability grouping, individualisation of programmes, team teaching, use of audio-visual aids.

Much of the research reported so far in this section is classroom focussed but what of factors that affect learning from outside the classroom itself? Several research studies are reported. Sammons, Hillman, and Mortimore (1995) review recent literature on school effectiveness. The writers acknowledge that there is a considerable body of evidence which shows that individual student background characteristics account for a larger proportion of total variance than does the school attended; nonetheless, there is also evidence to show that schools can make a significant difference.

There seem to be a number of features common to effective schools. These include professional leadership, shared vision and goals, an effective learning environment, concentration on teaching and learning, purposeful teaching, high expectations, positive reinforcement, monitoring progress, pupils rights respected, a home school partnership and effective professional development. The authors point out that classroom practice is the key element in effectiveness. Effective schools are likely to be calm, task oriented and orderly. School effectiveness is dependent on effective classroom teaching. This is teaching which maximises learning time, emphasises academic content and focuses on achievement. Effective teaching is purposeful, it has clear attainable goals and clarity of purpose. Lessons are structured sequentially, make clear what has to be learned and break teaching sequences into manageable segments. Effective teachers use student prompts and ideas and regularly test for progress.

Newmann and Wehlage (1998) synthesises 5 years of research conducted by the Centre on Organisation and Restructuring of Schools (CORS). The research involved several thousand students in hundreds of schools. A conclusion is that the recent education reform movement gives too much attention to changes in school organisation that do not directly address the quality of student learning. Student learning can meet high standards if educators and the public give students three kinds of support - teachers who practise authentic pedagogy, schools that strengthen professional community, and supportive external agencies and parents.

Wang, Haertel and Walberg (1993) conducted an extensive review of research to identify and estimate the influence of educational, psychological, and social factors on learning. They used evidence from three sources: 61 research experts, 91 meta-analyses, and 79 handbook chapters and

narrative reviews. The reviewers found moderate to substantial agreement about the findings from the large number of empirical studies, across three methods of analysis. First, distal variables (such as state district, and school level policy, and demographics) had little impact upon school learning. The reviewers note that distal variables do not impact directly upon most students' learning, since students in a particular classroom may have a teacher with a different interpretation of policy.

Second proximal variables strongly influenced school learning. They categorise these as

- psychological aptitudes (such as metacognitive, cognitive, motivational, and affective variables), noting that psychological theorists have come to view learners as 'architects' who build their own knowledge structures, and
- instructional variables (classroom management techniques, and the quality of the academic interactions between teacher and students).

The home environment provides a third set of variables. The home functions as an important site that supports students learning through parents' interest and involvement in student learning, and their expectations. The home is described as the most important out-of-school variable.

Another issue is whether the mix of students in a school affects learning.

Thrupp (1997) carried out an empirical study of the affect of the mix of a school's student population upon student achievement. In a qualitative study of four secondary schools, over 260 lessons were observed. It was concluded that learning was related to the social class of the students, and in the schools there was evidence that a relationship existed between social class and teaching methods, assessment procedures, teacher qualifications, and curriculum depth covered in lessons. These effects were not quantified, so it is not possible to assign a precise level of effect to any variable.

Other research has explored the cultural dimension of student learning and social outcomes (Chapple et al, 1997; Clark, 1998; Mitchell and Brumfit, 1997; Ogbu, 1995, Irvine and York, 1995; Wlodkowski and Ginsberg, 1995). Their commentary on attempts to enhance achievement notes the complexity of this issue, and the need to see the individual student as the focus for enhancement rather than taking a generalised approach. The suggestions for intervention map well onto recommendations such as those of Black and Wiliam (1998) about the value of formative assessment, and Hattie (1999), being to have high expectations of all students, and that teachers need to have knowledge of culture which can inform interactions with individual students. Chapple et al (1997) note the paucity of research in New Zealand to inform the many issues related to Maori education (see Chapter 8).

Some studies looked to establish that an aspect of schooling makes a difference in raising achievement and eventual employment prospects for students from different cultures. For example, Hunter (199x) looks at schooling in general. The report concluded that education is the largest single factor associated with the current poor outcomes for indigenous employment in Australia. In fact the influence of education dwarfs the influence of most demographic, geographic and social variables.

Collins (1994) indicates the complexity of enhancing educational outcomes through compensatory programmes for disadvantaged children on the basis of economic factors. The author points out that any programme designed to bring about greater equality by addressing inequalities on the basis economic factors will be controversial and contested. He notes that disadvantaged groups do not form a single group, being divided internally according to a number of factors. These divisions provide a complex set of viewpoints from within the general group about the solution to the problems of inequality. Factors such as racism, regionalism and class become important in determining sub groups' aspirations.

The author argues that solutions to problems of inequity on the basis of poverty are still a long way off. The long term solution to the problems involves social alliances whose outlines are still just emerging.

There are issues regarding bilingual learners that impact upon learning. Mitchell and Brumfit (1997) report a study of the English and foreign languages classroom experiences of year 9 pupils in England and Wales. In this project, researchers worked in three schools, one of which had significant numbers of bilingual learners, and observed/recorded lessons, interviewed teachers and pupils and recorded pupils performing a number of language related tasks. The main focus of the project was to explore the intent and nature of the pupils' and teachers' 'knowledge about language' (KAL), how KAL issues were handled in the classroom and possible links between KAL and pupils' language development. The paper compares the experience and responses of bilingual learners in mainstream classrooms with those of monolingual English speakers, using factors such as awareness of language variation and language learning strategies. The research account contains a critical discussion of policy developments for language education and an exploration of the meaning of 'equal opportunity' for bilingual pupils. The authors argue that the English-medium National Curriculum in the school with significant numbers of bilingual/multilingual students was clearly tailored to mainstream content. That is, the KAL units of work on language variation were not taught on bilingualism/multilingualism but on variation within English. No community or home language was available in the National Curriculum 'modern foreign languages' classes, but this was provided it as an after school activity.

Fordham (1998) reports two studies in which the use of indigenous language in the US is claimed to have had a positive effect on achievement in standardised tests, and in language acquisition.

West, Noden and Hedge (1998) contributes a different perspective to barriers and enhancements to educational achievements by studying the impact of parents on school achievement in their children. The main finding is that there is a very strong link between the mother's educational achievement and the educational achievement of the child. It was noted that social class seems less important than mother's educational level. The authors recommend further research on this issue.

The interaction between assessment and pedagogy, and implications for learning outcomes is a theme in the next group of research reports. Newton, Driver and Osborne (1999) explored the external pressures imposed upon science teachers by the National Curriculum in England and Wales and its assessment system. Detailed observations of 34 lessons revealed that limited pedagogical approaches were taken by teachers. Science lessons became teacher dominated with heavy emphasis on teacher exposition. Time pressure imposed by the need to cover the National curriculum was seen as a powerful factor mitigating against the greater involvement of students in knowledge construction.

Silvernail (1996) interviewed 50 teachers, head teachers, educational officials and policy researchers in 1993 in England and Wales when the National Curriculum had been introduced. He found that the majority of teachers had fundamentally changed the way they taught, what they taught and how they assessed their students. Whereas the intended curriculum was to provide each child with a broad and balanced curriculum, it appeared that this did not happen, even though teachers made considerable efforts to implement it. One problem was that curricula were built piecemeal without reference to other curricula. The amount of instructional time was found to relate directly to the national testing. The national curriculum was seen by teachers to be diluting the classroom curriculum by limiting the time available to explore new areas.

The prescribed assessment program also contributed to narrowing the curriculum - with the assessed curriculum becoming the taught curriculum. The impact of the SATs (pen and paper version) meant that teachers felt they should separate teaching from assessing. With the introduction of national testing and league tables temporary and permanent exclusions increased 66%. In their discussion of the impact of national assessment the authors cite three studies which have documented how high stakes traditional tests resulted in narrowing the curriculum, and encouraged teaching to the test. However, they argue, low stakes assessment (with elements of formative assessment) have greater potential for increasing performance.

Hilton (1998) contributes to the concerns about the influence of external policy on pedagogy and learning when commenting on the "Literacy Hour" a compulsory daily one hour programme of literacy instruction introduced into British schools in 1998. This hour is made up of instruction and set group tasks. Hilton points out that research has shown that using one approach to literacy education is less effective than using a variety of approaches, and that research shows that the type of direct instruction advocated in the Literacy Hour to be ineffectual. "The long tail" of pupils achieving low scores in international literacy tests is further argued to be the result of social and economic factors. Accurate testing of literacy achievement requires longitudinal studies which have not been carried out for programmes involving more phonics or direct instruction. The author argues that British literacy policy should be concerned with mixing socially advantaged and disadvantaged students, and with establishing systematic encouragement of parental interest and support.

Darling-Hammond (1994) develops the theme of assessment from a different perspective, that of social equity and implied social outcomes. She argues that the uses of assessment must be rethought, since the goal of schooling is now stated to be to educate all children well, rather than preparing some for an academic curriculum. Inequalities in access to education need to be tackled directly if all students are to be well educated. Testing students will not provide accountability in education while some students receive only a fraction of the school resources that support their more privileged counterparts. For all students to receive high quality instruction from highly qualified teachers, financial investments in schooling should be equalised across rich and poor communities. The author argues that if schools are financially supported, authentic assessment strategies can help schools become educational communities committed to self determined common values. Assessment may then work on behalf of equity in education, rather than perpetuating the kinds of inequalities that now exist.

Summary

This considerable literature is broadly coherent with the analytical framework for the review (Figure 2). A complex of interactions between curriculum, assessment and pedagogy affects students' learning and social outcomes. The central role of the teacher is particularly apparent when all these reports are assembled. There is sufficient commentary in the previous section to provide an overview of the specific questions posed for this review.

Specific Questions to Guide the Review

The review focus questions seek research-based answers to the following relationships between curriculum, assessment and pedagogy and their learning and social outcomes.

- 1. The effect of curricula on pedagogical approaches and students' learning and social outcomes.**
 - According to the research, what are the most significant effects of different kinds of curriculum specification (for example, outcomes based) on :
 - pedagogical approaches; and on
 - students' learning and social outcomes, particularly those for Maori and Pacific students, students from low decile and immigrant backgrounds and students of differing abilities as well as the learning and social outcomes of boys as compared with girls?
 - What conclusions can be drawn from research which compares the effectiveness of pedagogical approaches in systems with a mandatory curriculum and those with other curricula arrangements?
 - What new light can be shed on established pedagogical beliefs and approaches in the classroom in relation to curricula policies and to learning and social outcomes.

- 2. The effect of assessment regimes on pedagogical approaches and students' learning and social outcomes?**

- What does the research say about the extent to which the presence or absence of particular assessment regimes influences students' learning and social outcomes?
- What are the most significant effects of assessment regimes on pedagogical approaches reported in the literature?
- What does the literature say about the effects of assessment regimes on teaching practice at different levels of schooling, for example, primary and secondary levels?
- What does the literature say about the effect of different kinds of assessment regimes on the student groups identified above?

This literature review found no evidence for direct influences of the kind implicit in the first questions about curriculum, since there are strong intervening influences on curriculum from assessment procedures and pedagogical approaches. The specific details of the pedagogical approaches actually used in classrooms in the many studies surveyed were difficult to ascertain. This meant that conclusions about the interplay between curriculum, pedagogy and learning and social outcomes are difficult to quantify. Yet these interactions are very important. The research literature surveyed also lacked robust analysis of the influence of different curriculum specification on different ethnic groups, students from low decile backgrounds, and on boys' as against girls' achievement.

The major comparative studies of Black and colleagues note that OECD countries with substantially different curriculum specifications do not cluster significantly in their students' results in international surveys. The new curricula in England and Wales provided a structure for teaching which was initially welcomed as a clarification of the requirements for classroom learning, and teachers worked together to develop pedagogical approaches to implement these. However the associated assessment regimes created some resistance when teachers felt that they interfered with best practice. The greater the specification of curriculum, and the higher the stakes associated with linked assessment processes with a heavy emphasis on summative processes, the more it is likely that restricted pedagogical approaches with a focus on fragmented and rote learning will occur.

From the literature surveyed the most helpful response about curriculum would appear to be that there is no certainty about whether learning achievements can follow when there is specification of sufficient detail to help teachers guide student learning, with clear indications of how subject material should be structured; as opposed to less structured curriculum. There may be problems associated with finely-detailed curriculum specification of content when this can lead to over-emphasis on teaching for recall of detail rather than for understanding and connecting relevant information.

A response to the second set of questions on assessment regimes became somewhat clearer after the literature review. Major reviews of formative assessment provide indications of some learning gains. The reviews show that learning may be improved when teachers work with learners to integrate testing and instruction, with a clear conception of the curriculum, and of the goals and the process of instruction. Such a process has a role for summative assessment when this is introduced in such a manner as to provide for feedback and remediation. Little literature was reported on issues related to race, class and gender, but we are aware of a large American literature which was not reviewed. This gap means that no robust answers can be given to the questions about different groups in the classroom.

With respect to pedagogy, it is valuable to note, in the context of learning and social outcomes, that there is broad agreement in the literature surveyed that teaching approaches which begin with high expectations of students, contain high levels of work-related talk and discussion, and build from students' existing knowledge contribute positively to learning enhancement. It appears that learning outcomes are enhanced by teachers having an understanding of conceptual and procedural development in the related subject area.

Chapter 8: Gaps and recommendations for Research

The review has surveyed research into the complex processes of teaching and learning to enhance student achievement. Throughout this analysis there has been reference to the need for greater clarity in the research focus on the issues involved. Teaching and learning in the classroom provides a very complex research arena; the comments about insufficient clarity in research stems from the many interacting variables affecting change in this system. At the classroom level, which is the focus of learning, there is often insufficient detail about key features such as:

- how did teachers interpret the official curriculum?
- what assessment processes were actually implemented?
- what pedagogical approaches applied in the particular classroom being investigated?

The review found many large analyses which had sought general factors which would lead to learning enhancement, and described some which look significant; formative assessment, addressing the needs of individual students through high expectations, supportive and clearly framed learning experiences, and having teachers confident in their knowledge of the structures and main ideas of the subject they are teaching. Issues to do with different cultures, classes and genders were disappointingly invisible in the articles reviewed. The first consideration is of particular relevance to New Zealand education.

There are, then, considerable knowledge and research gaps. The compilers of this analysis find the comments in Black and Wiliam (1998) important. In their major review of assessment and classroom learning they state that research studies differed in the nature of the data which had been collected - or ignored, and that some determining features were often given no attention. They list research prospects and needs, some of which are relevant here. Important features requiring clarification include:

- issues relating to race, class and gender;
- the nature of the social setting in the classroom, as created by the learning and teaching members and by the constraints of the wider school system as they perceive and evaluate them;
- the perceptions and beliefs of teachers about learning, and about the abilities and prospects for their students;
- the assumptions about learning underlying curriculum and pedagogy;
- the precise nature of the various types of evidence revealed by student responses;
- the learning work used in acting on the interpretations;
- the divisions of responsibility between learners and teachers.

Black and Wiliam (1998) recognise that to adequately report on all of these aspects would be very difficult, and state that there is clearly a need for richer qualitative studies of processes and interactions within the classroom to inform our understanding. This enrichment of the more generalised quantitative data seems crucial. Their comment about the confused picture of benefits to disadvantaged and low-attaining students (many studies in their review reported substantial gains, but others did not support them) is particularly worth consideration as this is one focus of this review's aims.

There are two major gaps in the research on the effects of different kinds of curriculum models and regimes upon student learning. There are different forms of national curriculum and they need to be explored. Does how curricula are presented make a difference to teaching and learning? What part do levels and students play? What kinds of teacher support to implement new curriculum are required? What forms of assessment are suitable for different forms of curricula?

In the following sections some key gaps are described under the headings of gaps associated with Maori education, gaps associated with Pacific Island Nations people and other groups, gaps in research on assessment, gaps in research on curriculum,

Gaps associated with Maori education

The first gap of particular relevance to New Zealand is the lack of attention to issues relating to race. Chapple et al (1997) have commented on how Maori education is an under-researched area. The gaps they identified are elaborated below.

- What has been the impact of Maori-initiated educational changes on Maori education performance and on labour market performance? Curriculum changes (Maori language curricula) and school organisation changes (kura kaupapa Maori and Maori immersion classes) have been introduced but there is very little research into connections between these significant Maori education developments and achievement. This research will require longitudinal studies and appropriate resourcing. This gap needs urgent attention as the first kohanga reo graduates are now about to enter university. We have already lost the opportunity to research and monitor the first generation of students. What are the consequences for conceptual learning, language development, and assessment? What can we learn about second language pedagogy and language proficiency?
- What are the effects on Maori students of curriculum and assessment initiatives? In particular, assessment for culturally different groups is not on the international research agenda. Hence, one can only conclude that assessment research is seen as being secondary to the research on pedagogy and knowledge at this point in time, or that assessment researchers do not see it as a variable worth investigating.
- What are the causal factors underlying participation in early childhood education?
- What happens to Maori in primary and intermediate schooling? Research on school-based barriers and their effects is needed. How do school structures affect Maori students, particularly Maori students learning? Research is needed to explore effective teaching and learning which includes: teaching and learning styles; the cultural dimensions of learning and motivation; social and cultural factors that inhibit or enhance learning; the effects of role models and mentors on Maori students; exceptional and gifted Maori students.

Other key gaps requiring research are:

- Teacher change with respect to a socio-cultural framework. Whilst there is a large amount of literature looking at culturally appropriate strategies that may take account of the student learning, the role of the teacher has been largely untouched. Can classroom interactions be simply "managed" as a pedagogical moment requiring racial, cultural or gender sensitivity? Is familiarity with various cultural practices and/or cultural knowledge all that is required? What are the appropriate pre-service and in-service models and levels for producing bicultural/bilingual teachers?
- Cross-cultural interpersonal relationships. The area of postcolonial literature, exploring the role of the unconscious in how culturally different groups see each other and why, needs to be incorporated into educational research and explored as a possible reason for low expectations of Maori children by teachers.
- Research in education and te reo Maori. This is the an aspect of the international education debate that New Zealand could lead since research is vitally important worldwide for minority cultures trying to revitalise their language. Our kura kaupapa Maori are now well established and they should have had research and evaluation projects in them for every aspect of teaching and learning, including conceptual learning, language development, and assessment. In addition, language aspects such as second language pedagogy and language proficiency need investigation.

Gaps associated with Pacific Island Nations people and other groups

Our society is multicultural. Many of the research approaches suggested above may provide information helpful to other groups, but they also face different challenges. A programme of classroom based longitudinal research would contribute to better understanding of educational issues facing these students, since they would no longer be invisible in the wider picture.

Gaps in research on assessment

The aspect of teachers' assessment skills and procedures that is most in need of further research and development in terms of the findings from this literature review is the formative one. Research on what actually happens in classrooms (rather than reports about it from participants) is crucially needed. There is a lack of research in classrooms indicating the advantages and disadvantages of different forms of assessment tasks on student learning outcomes. No research findings on assessment activities which would provide valid, reliable, trustworthy and fair information about Maori, Pacific Nation or other minority students was found in the review. There is little if any research (particularly in New Zealand) on whether national testing and other accountability measures promote education change or development.

In the light of the concern that has been expressed about the quality and usefulness of data from large scale summative assessments (for example, IEA, TIMMS) in the development of national assessment policy there is need to develop reliable measures of actual classroom achievement which do not distort the learning process. We also need reliable measures of learning processes and how they impact upon learning and student understanding. This will contribute to knowledge about links between assessment and learning. Further research is also needed to assess the pros and cons of external testing and other summative measures.

Gaps in research on curriculum

More classroom research is needed on the interaction between curriculum, assessment and student achievement to develop a greater understanding of this complex area. Issues to be explored in relation to the New Zealand Curriculum Framework include the relationship between official curricula and the implemented classroom curricula, and the relationship between progressions defined in curricula and the actual learning progressions in the classroom. Classroom based research would also provide information on the way that the knowledge in the essential learning areas, the essential skills, and the attitudes and values defined in the curriculum are being addressed in New Zealand schools.

Further research is also needed on how curriculum documents influence teachers' classroom behaviour. This would assist in the development of policy directions regarding curriculum resources.

Curriculum documents also impact upon test items in national tests, but there seems to be relatively little connection between classroom usage of assessment items and curriculum documents. Further research may reveal relationships between curriculum, pedagogical approaches and assessment.

Gaps in research on pedagogy

The central role of the teacher in the classroom is clear from the research literature. Classroom based research which illuminates the particular importance of teachers: having a broad understanding of curriculum aims and objectives; having a wide range of pedagogical strategies; having high expectations of all students; knowing their students well; providing good feedback to students; recognising student success; having sound content knowledge of their subject; and taking responsibility for the progress of their students would contribute to teacher training and professional development programmes.

The importance of cultural awareness has already been noted as a significant research topic.

Major reform requires teacher professional input, and teacher development needs to parallel the reform. The research literature on teacher professional development contains gaps related to a number of deep and personal factors associated with pedagogical practice which may require change.

John Olson, one of the international consultants focused on innovations and the need to understand the factors influencing their uptake and dissemination. He provided the following research recommendations.

Research Issue 1

Reform entails some strategy of dissemination but little is known about how and if effective practitioners encourage others to innovate.

What role do advanced teachers play in assisting others in developing innovative strategies?

How can curriculum planning make provision for effective peer interaction?

Research Issue 2

Current innovation strategies are based on adequate understanding of the norms dominant in teacher and student culture, how non-engaged teachers compare to those who become engaged, and how disengagement from those norms can be supported.

How can teachers achieve the needed capability to engage in innovation in spite of norms which might lead to reluctance?

Research Issue 3

Existing change models tend to focus on management issues rather than the culture of practice. The problem of assimilating new technology is a cultural one, and may require an anthropological approach. There is a need for a balanced approach to research methodology.

What provision is there for intensive school-based research to be undertaken? Are school systems well equipped? How valid is the idea of the teacher researcher and action research?

Research Issue 4

Research is needed to find out more about teachers' work culture, the technologies that sustain it, and the implications of new approaches for those technologies. A set of research issues emerge:

- How well do curriculum innovations work in practice and how well are teachers able to implement associated constructivist ideas in practice?
- What role might teachers play in setting research agendas?
- How is feedback from classroom experience to be used? What role will different players have: researchers, administrators, policy makers? What forums need to be established to allow these parties to discuss school-based research?

Olson then comments about the need for a research base for teacher professional development, including in-service contributions, noting that the teacher voice is currently not central, since educational change is highly politicised and governments are lobbied by powerful stakeholders. If these stakeholders are overly influential the language of discussion and modes of management result in a reduced framework for discussion and action. Curriculum policies may be mandated without adequate tests of their validity and reliability. Management theories can offer ideas to improve practice, but educational values and the contribution that classroom practice can make to reform should not be ignored.

Research methodology

This review has noted comments on the need for data to be collected on many variables which may affect learning outcomes in the classroom. A significant problem has been that individual differences in implemented curriculum and assessment processes have been invisible, and information about pedagogical approaches used has also been unavailable. Large scale quantitative research frequently attempts to illuminate the effects of these variables through variable analysis, but this is difficult in the absence of relevant data. A more appropriate research process would be detailed classroom studies using explicit research questions related to these variables. Longitudinal studies would greatly increase the value of research studies by tracking learning outcomes when substantial changes in classroom procedures are investigated.

Another aspect of pedagogy that urgently needs researching is the learning styles of students and how these are linked to teaching approaches. A good deal has been about this linkage, but there is little research (especially New Zealand) on this issue.

A further research need is for studies of classrooms 'at work'. Little is known how New Zealand secondary school classrooms function and what students take from their experiences in them, and not much more about primary classrooms. Such studies would reveal how curriculum and assessment procedures shape students' learning, and how effective the time set aside for learning actually is. Information of this kind should reveal why some students and student groups fail.

References

- Adey, P. (1997, March 24 - 28). *Factors influencing uptake of a large scale curriculum innovation*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Alexander, R. (1995). *Versions of primary education*. London: Routledge.
- Alexander, R., Wilcocks, J., & Nelson, N. (1995). Discourse, pedagogy and the national curriculum : Change and continuity in primary schools. *Research Papers in Education, 11*(1), 81-120.
- Allen, N. (1997, March 21 - 24). *Indigenous models for science and culture curriculum development*. Paper presented at the National Association for Research in Science Teaching, Oak Brook, Illinois.
- Alton-Lee, A., & Nuthall, G. (1992). A generative methodology for classroom research. *Education Philosophy and Theory, 24*, 29-55.
- Alton-Lee, A., Nuthall, G., & Patrick, J. (1993). Reframing classroom research: a lesson from the private world of children. *Harvard Educational Review, 63*(1), 50-84.
- Anderson, G. L. (1998). Towards authentic participation: Deconstructing the discourses of participatory reform in education. *American Educational Research Journal, 35*(4), 571-606.
- Atkin, J. M. (1998). The OECD study of innovations in science, mathematics and technology education. *Journal of Curriculum Studies, 30*(6), 647-660.
- Bage, G., Grosvenor, J., & Williams, M. (1999). Curriculum planning: Prediction or response? A case study of teacher planning conducted through partnership action research. *The Curriculum Journal, 10*(1), 49 - 70.
- Becker, W. E. (1998). Standards and testing : Another view. *Journal of Economic Education, 29*(2), 183 - 186.
- Bell, B. (1993). *I Know about LISP but how do I put it into practice?* . Hamilton: University of Waikato.
- Bell, B., & Cowie, B. (1997). *Formative assessment and science education* . Hamilton: University of Waikato.
- Bell, B., Jones, A., & Carr, M. (1995). The development of the recent national New Zealand science curriculum. *Studies in Science Education, 26*, 73-105.
- Black, P. (1997). Can teachers use assessment to improve learning? *British Journal of Curriculum and Assessment*.
- Black, P., & Atkin, M. (Eds.). (1996). *Changing the Subject: Innovations in Science, Mathematics and Technology Education*. London: Routledge.
- Black, P., & Atkin, M. (1997). Policy perils of international comparisons: The TIMSS case. *Phi Delta Kappan, 79*(1), 22-28.
- Black, P., Brown, M., Simon, S., & Blondel, I. (Eds.). (1996). *Progression in learning: Issues and evidence in mathematics and science*. Oxford, United Kingdom: Blackwell Publishers.
- Black, P., & Michel, A. (Eds.). (1998). *Learning from pupil assessment: International comparisons*. Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education and Information Studies.

- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, Policy and Practice*, 5(1), 7-74.
- Bracey, G. W. (1996). International comparisons and the condition of American education. *Educational Researcher*, 25(1), 5-11.
- Broadfoot, P. (1998, April 13-17). *Categories, standards and instrumentalism: Theorising the changing discourse of assessment policy in English primary education*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Brophy, J., & Alleman, J. (1992). Planning and managing learning activities. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 3, pp. 1-45). Connecticut: JAI Press.
- Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching*. New York: Macmillan.
- Brown, M. (1996). FIMS and SIMS: the first two IEA international mathematics surveys. *Assessment in Education*, 3(2), 193-212.
- Brown, M., Taggart, B., McCallum, B., & Gipps, C. (1996). National testing: the impact of key stage 2 tests. *Education 3 to 13*, 24(3), 3-7.
- Carlgren, I. (1999). Professionalism and teachers as designers. *Journal of Curriculum Studies*, 31(1), 43-56.
- Case, R. (1991). *The mind's staircase*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Centre for Civic Education. (1994). *National standards for civics and government*. Calabasas, California: Centre for Civic Education.
- Centre for Civic Education. (1998). *Report of the task force on civic education*. Washington: Centre for Civic Education.
- Chapple, S., Jefferies, R., & Walker, R. (1997). *Maori participation and performance in education: a literature review and research programme*. Wellington: New Zealand Institute of Economic Research.
- Cheng, L. (1999). Changing assessment: Washback on teachers' perceptions and actions. *Teaching and Teacher Education*, 15(3), 253 - 271.
- Clark, M. (1998). *Maori and Pacific Island performance in mathematics*. Paper presented at the Mathematics Education Research Seminar, Wellington, New Zealand.
- Clarke, S. (1996). The impact of national curriculum statutory testing at key stages 1 and 2 on teaching and learning and the curriculum. *British Journal of Curriculum and Assessment*, 7(1), 12-18.
- Cogan, J. J. (1997). *The citizenship education policy study project. Final report*. Canberra, Australia: Citizen Education Policy Study Project.
- Cohen, E. G. (1994). Restructuring the classroom. *Review of Educational Research*, 64(1), 1-35.
- Collins, R. W. (1994). Poverty and education. *Harvard Educational Review*, 64(2), 125 - 149.
- Corcoran, T., & Goertz, M. (1995). Instructional capacity and high performance schools. *Educational Researcher*, 24(9), 27-31.

- Cowie, B., & Bell, B. (1995). *Views of assessment. Research report 1, learning in science project (assessment)*. Hamilton: University of Waikato.
- Crombie, W. H., & Johnson, D. E. (1998). Assessment for success in Maori-medium primary schooling? *New Zealand Journal of Educational Studies*, 33(2), 143-153.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Darling-Hammond, L. (1994a). National standards and assessment: Will they improve education? *American Journal of Education*, 102(No.4), 478 - 510.
- Darling-Hammond, L. (1994b). Performance based assessment and educational equity. *Harvard Educational Review*, 64(1), 5 - 30.
- Darling-Hammond, L., & Falk, B. (1997). Using standards and assessments to support student learning. *Phi Delta Kappan*, 79(3), 190-199.
- Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives*, 6(1).
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53(2), 159-199.
- Doyle, W. (1992). Curriculum and pedagogy. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 486-516). New York: Macmillan.
- Doyle, W., & Carter, K. (1984). Academic tasks in the classroom. *Curriculum Inquiry*, 14(2), 129-149.
- Dunkin, M. J., & Welch, A. (1996, July). *Four case studies of teacher knowledge and citizenship education*. Paper presented at the Australian Teacher Education Association.
- Education Review Office. (1999). *Student assessment: Practices in primary schools* (2). Wellington, NZ: Education Review Office.
- Eisner, E. (1999). The uses and limits of performance assessment. *Phi Delta Kappan*, 80(9), 658-660.
- Eisner, E. W. (1994). *The educational imagination*. New York: Macmillan.
- Elley, W. (Ed.). (1994). *The IEA study of reading literacy: achievement and instruction in thirty-two school systems* (Vol. 11). Oxford: Elsevier Science.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: MacMillan.
- Evertson, C. M., & Green, J. L. (1986). Observation as inquiry and method. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 162-213). New York: Macmillan.
- Fearnely-Sander, M., & Sprot, T. (1996). *The pedagogy of civics and citizenship education*. Melbourne: Australian Curriculum Corporation.
- Fordham, M. (1998). The politics of language and the survival of indigenous culture: From suppression to reintroduction in the formal classroom. *Equity and Excellence in Education*, 31(1), 40-47.
- Fore, L. (1998). Curriculum control: Using discourse and structure to manage educational reform. *Journal of Curriculum Studies*, 30(5), 559-576.

- Friedman, V. (1997). Making schools safe for uncertainty: Teams, teaching and school reform. *Teachers College Record*, 99(2), 335-370.
- Friesen, J. W. (1996, September 27 - 28). *The concept of giftedness in the context of first nations culture and philosophy*. Paper presented at the Faces of Excellence, University of Calgary, Calgary, Alberta, Canada.
- Galton, M., Hargreaves, L., & Comber, C. (1998). Classroom practice and the national curriculum in small rural primary schools. *British Educational Research Journal*, 24(1), 43 - 61.
- Galton, M., Hargreaves, L., Comber, C., Wall, D., & Tell, T. (1999). Changes in patterns of teacher interaction in primary classrooms: 1976 to 1996. *British Educational Research Journal*, 25(1), 23 - 37.
- Gamoran, A. (1997). Curriculum change as a reform strategy: Lessons from the US and Scotland. *Teachers College Record*, 98(4), 608-628.
- Gamoran, A., & Weinstein, M. (1998). Differentiation and opportunity in restructured schools. *American Journal of Education*, 106(3), 385 - 415.
- Garbutcheon Singh, M. (1996). *Citizenship in Australia's multicultural society*. Melbourne: Curriculum Corporation of Australia.
- Garden, R. A. (Ed.). (1996). *Science performance of New Zealand Form 2 and Form 3 students: National results from New Zealand's participation in the third international mathematics and science study*. Wellington: Research and International Section Ministry of Education.
- Garden, R. A. (Ed.). (1997). *Mathematics and science performance in middle primary school: Results from New Zealand's participation in the third international mathematics and science study*. Wellington: Research and International Section Ministry of Education.
- Gay, G. (1997). The relationship between multicultural and democratic education. *The Social Studies*, 88(1), 5 -10.
- Gilbert, R., & Gilbert, P. (1998). *Masculinity goes to school*. St. Leonards, Australia: Allen & Unwin.
- Gipps, C. (1996a, April 13 - 17). *Assessment for the millennium: Form, function and feedback*. Paper presented at the Inaugural lecture, Institute of Education, University of London.
- Gipps, C. (1996b, September 20-25). *Findings from large scale assessment in England*. Paper presented at the 22nd Annual IAEA Conference, Beijing, China.
- Gipps, C. (1999). Sociocultural aspects of assessment. *Review of Research in Education*, 24.
- Gipps, C., Clarke, S., & McCallum, B. (1998). *The role of teachers in national assessment in England*. Paper presented at the American Educational Research Association, San Diego.
- Gipps, C., & MacGilchrist, B. (1999). Primary school learners. In P. Mortimore & P. Chapman (Eds.), *Pedagogy and its impact on learning*.
- Gipps, C., McCallum, B., & Brown, M. (1999). Primary teachers' beliefs about teaching and learning. *The Curriculum Journal*, 10(1), 123-134.
- Gipps, C., & Tunstall, P. (1998). Effort, ability and the teacher: Young children's explanations for success and failure. *Oxford Review of Education*, 24(2), 149-165.

- Gitlin, A., & Margonis, F. (1995). The political aspect of reform: Teaching resistance as good sense. *American Journal of Education.*, 103.(No.4.), 377 - 405.
- Goldstein, H., & Blatchford, P. (1998). Class size and education achievement: A review of methodology with a particular reference to study design. *British Educational Research Journal*, 24(3), 255 - 268.
- Greenwald, R., Hedges, L. U., & Laine, R. D. (1996). The affect of school resources on student achievement. *Review of Educational Research*, 66(3), 361-396.
- Griffin, P. (1998). Outcomes and profiles: Changes in teachers' assessment practices. *Curriculum Perspectives*, 18(1), 9 - 20.
- Haertel, E. (1999). Performance assessment and educational reform. *Phi Delta Kappan*, 80(9), 662-666.
- Haigh, N., & Katterns, R. (1984). Teacher effectiveness: problem or goal for teacher education. *Journal of Teacher Education*, 35(5), 23-27.
- Hall, K., & Nuttall, W. (1999). Of relative importance of class size to infant teachers in England. *British Educational Research Journal*, 25(2), 245 - 258.
- Hancock, L., & Kilpatrick, J. (1993). Effects of mandated testing on instruction. In N. R. C. Mathematical Sciences Education Board (Ed.), *Measuring what counts* (pp. 149-173). Washington DC.
- Hansen, D. (1999). Conceptions of teaching and their consequences. In M. Lang, J. Olson, H. Hansen, & W. Bunder (Eds.), *Changing schools/changing practices: Perspectives on educational reform and teacher professionalism* . Leuven: Grant.
- Harlen, W. (1999). *Effective Teaching of Science: A Review of Research*. Edinburgh: The Scottish Council for Research in Education.
- Harlen, W., Holroyd, C., & Byrne, M. (1995a). *Confidence and understanding in teaching science and technology in primary schools* . Edinburgh: Scottish Council for Research in Education.
- Harlen, W., & James, M. (1996, April). *Creating a positive impact of assessment on learning*. Paper presented at the AERA annual conference, New York.
- Harlen, W., Malcolm, H. C., & Bryne, M. (1995b). Teachers' assessment and national testing in primary schools in Scotland: Roles and relationships. *Assessment in Education*, 2(2), 129-144.
- Harris, B., & Drake, S. (1997). Implementing high school reform through school-wide action research teams: A three year case study. *Action in Teacher Education*, 14(3), 15-31.
- Hatch, T. (1998). The differences in theory that matter in school improvement. *American Educational Research Journal*, 35(1), 33 - 52.
- Hattie, J. (1999, August 20, 1999). The truth about better learning. *New Zealand Education Review*, 9.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: a meta-analysis. *Review of Educational Research*, 66(2), 99-136.
- Hill, J., & Hawk, K. (1998). Aiming for student achievement. *SET*, 2(4), 1-4.
- Hilton, M. (1998). Raising literacy standards : The true story. *English in Education*, 32(3), 4 - 16.
- Hoyle, E., & John, P. D. (1995). *Professional knowledge and professional practice*. London: Cassell.

- Hughes, P. (1994). *International best practice in civics education* . Australia: Civics Experts Group.
- Hunter, B. (unknown). The Determinants of Indigenous Employment Outcomes: The Importance of Education and Training. *The Australian National University*.
- Irvine, J. J., & York, D. E. (1995). Learning styles and culturally diverse students : A literature review. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of Research on Multicultural Education* (pp. 484-497). New York: Macmillan.
- Jackson, P. H. (1999). Teaching as a moral enterprise. In M. Lang, J. Olson, H. Hansen, & W. Bunder (Eds.), *Changing schools/changing practices: Perspectives on educational reform and teacher professionalism* . Leuven: Garant.
- James, M., & Gipps, C. (1998). Broadening the basis of assessment to prevent the narrowing of learning. *The Curriculum Journal*, 9(3), 285-297.
- Jones, K., & Whitford, B. L. (1997). Kentucky's conflicting reform principles. *Phi Delta Kappan*, 79(4), 276-281.
- Jones, L. (1996). *Reading national assessment*. Paper presented at the World literacy conference, Philadelphia.
- Joyce, B., Weil, M., & Showers, B. (1992). *Models of teaching* (4 ed.). Englewood Cliffs, New Jersey: Prentice-Hall.
- Kaltsounis, T. (1997). Multicultural education and citizenship education at a crossroads : Searching for common ground. *The Social Studies*, 18 - 21.
- Kellaghan, T. (1996). IEA studies and educational policy. *Assessment in Education*, 3(2), 143-160.
- Kennedy, K. J. (Ed.). (1996). *Conflicting conceptions of citizenship and their relevance for the school curriculum*. In *civics and citizenship education : Issues from practice and research*. Canberra: Australian Curriculum Studies Association.
- Keown, P. (1996). *Walking the sufficiency ridge* . Hamilton: University of Waikato.
- Klette, K. (1997). Teacher individuality, teacher collaboration and repertoire building : Some principle dilemmas. *Teachers and Teaching: Theory and Practice*, 3(2), 243 - 256.
- Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown Publishers.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465-491.
- Lawrence, S. M., & Tatum, B. D. (1997). Teachers in transition: The impact of antiracist professional development on classroom practice. *Teachers College Record*, 99(1), 162-178.
- Lee, V. E., & Smith, J. B. (1996). Collective responsibility for manning and its effects on games in achievement for early secondary school students. *American Journal of Education*, 104(2), 103 - 147.
- Levine, D. U., & Lezotte, L. W. (1995). Effective schools research. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of Research on Multicultural Education* (pp. 525-547). New York: Macmillan.
- Lewis, A. (1996). Summative national curriculum assessments of primary aged children with special needs. *British Journal of Special Education*, 23(1), 9-14.

- Linn, R. L. (1986). Quantitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 92-118). New York: MacMillan.
- Lohrey, A. (1996). *The pedagogy of civics and citizenship education : A review paper* . Melbourne: Curriculum Corporation of Australia.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: a meta-analysis. *Review of Educational Research*, 66(4), 423-458.
- Luke, C., & Gore, J. (Eds.). (1992). *Feminisms and critical pedagogy*. New York: Routledge.
- Madaus, G. (1988). *The influence of testing on the Curriculum*. Chicago: National Society for the Study of Education.
- Madaus, G. F. (1991). The effects of important tests on students. *Phi Delta Kappan*, 73(Nov), 226-231.
- Madaus, G. F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of Research on Curriculum* (pp. 119-156). New York: Macmillan.
- McCarthy, C., & Crichlow, W. (Eds.). (1993). *Race, identity and representation in education*. New York: Routledge.
- McKinley, E. (1995). *A power/knowledge nexus: Writing a science curriculum in Maori*. University of Waikato, Hamilton.
- Mehan, H., Okamoto, D., Lintz, A., & Wills, J. S. (1995). Ethnographic studies of multicultural education in classrooms and schools. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of Research on Multicultural Education* (pp. 129-144). New York: Macmillan.
- Mellor, S., & Elliot, M. (1996). *The pedagogy of civics and citizenship education* . Melbourne: Curriculum Corporation of Australia.
- Ministry of Education. (1993). *The New Zealand curriculum framework. Te anga matauranga o Aotearoa*. Wellington: Learning Media.
- Ministry of Education. (1999). *Gender differences in achievement and participation in the compulsory school sector* . Wellington: Ministry of Education.
- Mitchell, R., & Brumfit, C. (1997). The national curriculum experience of bilingual pupils. *Educational Review*, 49(2), 159-180.
- Morris, G. (1995). *Risky business: Documenting the students perspective on secondary schooling in New Zealand* . Hamilton: The University of Waikato.
- Murphy, P. (1996). The IEA assessment of science achievement. *Assessment in Education*, 3(2), 213-232.
- Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education*, 104(4), 280 - 312.
- Newmann, F. M., & Wehlage, G. G. (1998). *Successful school restructuring : A report to the public and educators*. Washington DC: Centre on Organization and Restructuring Schools, Madison WI, Office of Educational Research and Improvement.

- Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation of the pedagogy of school science. *International Journal of Science Education*, 21, 553-576.
- Nuthall, G. (1997). Understanding student thinking and learning in the classroom. In B. K. Biddle, T. L. Good, & I. Goodson (Eds.), *The International Handbook of Teachers and Teaching*. Dordrecht: Kluwer Academic Publishers.
- Nuthall, G., & Alton-Lee, A. (1992). Understanding how children learn in classrooms. In M. Pressley, K. Harris, & J. Guthrie (Eds.), *Promoting academic competence and literacy in school*. San Diego: Academic Press.
- Oakes, J., Wells, A. S., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform. *Teachers College Record*, 98(3), 482-510.
- Ogbu, J. (1995). Understanding cultural diversity and learning. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of research on multicultural education* (pp. 582-593). New York: Macmillan.
- Olson, J. (1992). In search of the expert teacher. In J. Olson (Ed.), *Understanding teaching*. Buckingham: Open University Press.
- Olson, J., E. J., & Lang, M. (1998). Changing the subject: The challenge to teacher professionalism of innovation in OECD countries. *Journal of Curriculum Studies*, 31(1), 69 - 82.
- Osborne, J., & Simon, S. (1996). Primary science: Past and future directions. *Studies in Science Education*, 26, 99-147.
- Ostermann, M. (1998). The dominator: Classroom computer use by young primary school children. *Teachers and Curriculum*, 2, 50-53.
- Parker, W. C. (1997). Navigating the unity/Diversity tension in education for democracy. *The Social Studies*, 88(1), 12 - 17.
- Parry, L. J. (1998). Immigration and multiculturalism : Issues in Australian society and schools. *Social Education*, 62(7), 449 -453.
- Pearson, J., & Bell, B. (1993). *The teacher development that occurred* (Research report). Hamilton: University of Waikato.
- Pennell, J., & Firestone, W. (1996). Changing classroom practices: Matching program features with teacher characteristics and circumstances. *Teachers College Record*, 98(1), 46-79.
- Phillips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal*, 34(4), 633 - 662.
- Phuntsog, N. (1998). *The magic of culturally responsive pedagogy: In search of a genie's lamp in multicultural education*. Paper presented at the American Educational Research Association, San Diego, California.
- Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. Phye (Ed.), *Handbook of classroom assessment* (pp. 53-68). California: Academic Press.
- Plewis, I., & Veltman, M. (1996). *Where does all the time go? Changes in pupils' experiences in year 2 classrooms*. Oxford, UK: Blackwell.

Pollard, A., Broadfoot, P., Croll, P., Osborn, M., & Abbott, D. (1994). *Changing English primary Schools? A Cautionary Analysis*. Paper presented at the AERA conference, New Orleans.

Porter, A. C., & Brophy, J. E. (1988). Synthesis in research on good teaching: insights from the work of the Institute of Research on Teaching. *Educational Leadership*, 48(8), 74-85.

Poston, W. K. (1997). Pretending to know what we don't know : A constructivist look at current assessment findings in curriculum audits. *International Journal of Educational Reform*, 6(1).

Prawat, R. S. (1995). 1990s curriculum reform: Deja vu all over again? In J. Brophy (Ed.), *Advances of Research on Teaching* (Vol. 5, pp. 13-33). Connecticut: JAI Press.

Preece, P. F. W., & Skinner, N. C. (1999). The national assessment in science at key stage 3 in England and Wales and its impact on teaching and learning. *Assessment in Education: Principles, Policy and Practice*, 6(1), 11 - 25.

Print, M. (1996a). *Conceptual design for the civics and citizenship education schools curriculum materials project*. Melbourne: Curriculum Corporation of Australia.

Print, M. (1996b). *Pedagogical strategies for civics and citizenship education*. Melbourne: Curriculum Corporation of Australia.

Raizen, S., & Britton, E. (1997). *Bold ventures*. Dordrecht: Kluwer.

Reed, L. R. (1999). Troubling boys and disturbing discourses on masculinity and schooling: A feminist exploration of current debates and interventions concerning boys in school. *Gender and Education*, 11(1), 93-110.

Robertson, C., Cowell, B., & Olson, J. (1998). A case study of integration and destreaming: Teachers and students in an Ontario high school respond. *Journal of Curriculum Studies*, 30(6), 691-717.

Roelofs, E., & Terwel, J. (1997). *Constructivism and authentic pedagogy : State of the art and recent developments in the Dutch national curriculum in secondary education*. Paper presented at the American Educational Research Association, Chicago, Illinois.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2), 181-221.

Ross, J. A. (1995). Strategies for enhancing teachers beliefs in their effectiveness: Research on a school improvement hypothesis. *Teachers College Record*, 97(2), 227-251.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(5), 77-84.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools*. London: Office for Standards in Education (OFSTED).

Schon, D. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Scott, D. (1997). The primary science curriculum. In B. Bell & R. Baker (Eds.), *Developing the science curriculum in Aotearoa, New Zealand*. Auckland: Addison Wesley Longman.

Shepard, L. A. (1991). Will national testing improve student learning? *Phi Delta Kappan*, 73(November), 232-238.

- Shepard, L. A., Flexer, R. J., Hiebert, E. H., Marion, S. F., Mayfield, V., & Weston, T. J. (1996). Effects of introducing classroom performance assessment on student learning. *Educational Measurement: Issues and Practice*, 15, 7-18.
- Sheppard, L. A. (1995). Using assessment to improve learning. *Educational Leadership*(February 1995), 38-43.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.
- Silvernail, D. (1996). The impact of England's national curriculum and assessment system on classroom practice: Potential lessons for American reformers. *Educational Policy*, 10(1), 46-62.
- Sizmur, S., & Sainsbury, M. (1997). Criterion referencing and the meaning of national curriculum assessment. *British Journal of Educational Studies*, 45(No.2), 123 - 140.
- Slack, J. B. (1998, April 13 - 17). *A model for measuring mathematical achievement test performances : A longitudinal analysis of non traditional learners*. Paper presented at the American Educational Research Association, San Diego.
- Slavin, R. E. (1995). Co-operative learning and intergroup relations. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of Research on Multicultural Education* (pp. 628-634). New York: Macmillan.
- Sleeter, C. E. (Ed.). (1993). *How white teachers construct race*. New York: Routledge.
- Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 20(Winter), 7-11.
- Snively, G., & Corsiglia, J. (1998, April 19 - 22). *Discovering indigenous science : Implications for science education*. Paper presented at the National Association for Research in Science Teaching, San Diego, California.
- Snow-Renner, R. (1998, April 13 - 17). *Mathematics assessment practices in Colorado classrooms: Implications about variations in capacity and opportunity to learn*. Paper presented at the American Educational Research Association, San Diego.
- Spillane, J. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, 31(2), 143-175.
- Sternberg, R., & Horvath, J. (1995). A prototype view of expert teaching. *Educational Researcher*(August-September 1995), 9-17.
- Stodolsky, S. S., Salk, S., & Glaessner, B. (1991). Student views about learning math and social studies. *American Educational Research Journal*, 28(1), 89-116.
- Strand, S. (1997). Pupil progress during key stage 1: A value added analysis of school effects. *British Educational Research Journal*, 23(4), 471 - 487.
- Thrupp, M. (1997). How school mix shapes school processes: a comparative study of New Zealand schools. *New Zealand Journal of Educational Studies*, 32(1), 53-82.
- Timperley, H., Robinson, V., & Bullard, T. (1999). *Strengthening education in Mangere and Otara evaluation: first evaluation report (prepared for NZ Ministry of Education)* . Auckland: University of Auckland.

- Tunstall, P., & Gipps, C. (1996). 'How does your teacher help you to make your work better?' Children's understanding of formative assessment. *The Curriculum Journal*, 7(2), 185-203.
- Turner-Bisset, R. (1999). The knowledge basis of the expert teacher. *British Educational Research Journal*, 25(No.1.), 39 - 55.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, Mass: Harvard University Press.
- Veenman, S. (1995). Cognitive and none cognitive effects of multi grade and multi age classes: A best evidence synthesis. *Review of Educational Research*., 65(No.4), 319 - 381.
- Veenman, S. (1996). Effects of multigrade and multi-age classes reconsidered. *Review of Educational Research*, 66(3), 323-340.
- Waks, L. (1999). Reflective practice in the design studio and teacher education. *Journal of Curriculum Studies*, 31(3), 303-316.
- Walsley, P. (1999). Teaching worth celebrating. *Educational Leadership*(May 1999), 8-13.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Towards a knowledge base for school learning. *Review of Education Research*, 63(3), 249-294.
- West, A., Noden, P., & Hedge, A. (1998). Parental involvement in education in our schools. *British Educational Research Journal*, 24(4), 461 - 484.
- Whitty, G., Aggleton P, & Rowe, G. (1996). *Competing conceptions of quality and social education: Learning from the experience of a cross- curricular themes*. London: Blackwell.
- Whitty, G., Power, S., & Halpin, D. (1998). *Devolution and choice in education: The school, the state and the market*. Buckingham, United Kingdom: Open University Press.
- Wilson, J. (1996). *Assessment for student learning*. Paper presented at the Australia & New Zealand Conference on the First Years of Schooling, Hobart, Australia.
- Winklemann, R. (1998). *The economic benefits of schooling in New Zealand : Comment and update*. Christchurch, New Zealand: Department of Economics, University of Canterbury.
- Wlodkowski, R. J., & Ginsberg, M. B. (1995). A framework for culturally responsive teaching. *Educational Leadership*, 17-21.
- Wragg, E. C., Wragg, C. M., Hayes, G. S., & Chamberlin, R. P. (1998). *Improving literacy in the primary school*. London: Routledge.
- Yates, L. (1997). Gender equity and the boys debate: What sort of challenge is it? *British Journal of Sociology of Education*, 18, 337 - 347.
- Zepeda, S., & Pontecelli. (1998). At cross purposes: What do teachers need, want, get from supervision. *Journal of Curriculum and Supervision*, 14(1), 68 - 87.

Appendix One: Literature sources

• Literature sources

The criteria being used to select articles for reviewing and summarising were:

- the presence of material addressing the links and relationships between the nodes in the model (that is between curriculum, assessment, pedagogy, educational outcomes and social outcomes)
- major review articles within the nodes.
- meta-analyses of case studies within the nodes.
- articles by key researchers within the nodes.

• **Electronic sources.** To date, we have searched the electronic data bases of ERIC, Australian Education Index, British Education Index, Social Sciences Citation Index, Current Contents, Uncover, INNZ (Index New Zealand). The initial scans have presented over 2,000 abstracts and references to be considered for inclusion in the review. Of these 500 abstracts/references have been the focus of the developing report.

• Physical searches

Physical searches were done on the recommendation of Paul Black who noted the limitations of electronic searches. These limitations arise out of the varied definitions underlying the keywords used to organise the electronic data. Physical searches of the following journals for the last 5 years were carried out:

- Access
- American Education Journal
- American Journal of Education
- Assessment in Education
- British Educational Research Journal
- British Journal of Educational Psychology
- British Journal of Educational Studies
- Cambridge Journal of Education
- Canadian Social Studies
- Children's Social and Economic Education
- Contemporary Educational Psychology
- Delta
- Education Review
- Educational Evaluation and Policy Analysis (to 1979)
- Educational Leadership
- Educational Psychologist
- Educational Researcher
- Elementary School Journal
- Harvard Educational Review
- International Journal of Educational Research
- International Journal of Mathematics Education
- International Journal of Nursing Studies
- International Journal of Science Education
- International Journal of Technology and Design Education
- International Review of Education
- Journal for Research in Mathematics Education
- Journal of Curriculum Studies
- Journal of Educational Psychology
- Journal of Educational Research
- Journal of Research in Science Teaching

- Measurement Issues and Educational Practice
 - New Zealand Journal of Education Studies
 - Oxford Review of Education
 - Phi Delta Kappan
 - Pukenga
1. Research in Science Education
 2. Research Papers in Education
 3. Review of Educational Research
 4. Review of Research in Education
 5. Science Education
 6. Sites
 7. Social Education
 8. Studies in Science Education
 9. Teachers College Record
 10. Teachers and Teaching
 11. The Curriculum Journal
 12. The Social Studies
 13. Theory and Research in Social Education
 14. Waikato Journal of Education

Appendix Two: Consultants and search protocol

Research team structure and personnel

- **Core Project Team**

Co-directors

Professor Clive McGee

Dr Alister Jones

Team Members

Dr Hugh Barr

Associate Professor Beverley Bell

Elizabeth McKinley

Analyst/Writer

Dr Malcolm Carr

Information Professionals

Tina Simpson as research assistant

Kayrn Kee

- **Waikato/National Search and Guidance Team³**

Dr Ken Carr

Bronwen Cowie

Nola Campbell

Dr Mike Forret

Margaret Scratchley

George Salter

Clare Henderson

Dr Sue Dymock

Dr Terry Locke

Mary Hill

Ngaere Roberts

Pepi Walker

Harata Te Aika

Ngarewa Hawera

Angus MacFarlane

Cath Rau

³ Social Science and Science are covered by the core project team

• **International Search and Guidance Team**

Professor Paul Black	Emeritus Professor, Kings College London
Professor Caroline Gipps	Professor, Kingston University, London
Professor John Olson	Emeritus Professor, Queen's University, Ontario
Professor Olug Jegede	Professor, Open University, Hong Kong
Professor Brian Hill	Professor, Murdoch University, Western Australia

• **Reference Group**

Professor Noeline Alcorn	Education Policy
Professor Russell Bishop	Maori Education
Professor Ted Glynn	Special Education

Search protocol

Information sent to members of the national and international teams.

Ministry of Education ‘State-of-the-Art’ Literature Review

The Effects of Curricula and Assessment on Pedagogical Approaches and on Education Outcomes

- 1 A copy of the model (see **Figure 1 in body of the review**)
- 2 A copy of the primary and secondary keywords (see **Table 1 in body of the review**)
- 3 The following protocol

The Effects of Curricula and Assessment on Pedagogical Approaches and on Education Outcomes

PROTOCOL

NAME

FIELD(S)

- 1. What are the names of key journals and publications in the related areas eg. research reports, books, government reports**
- 2. What are the key names of people writing in the area.**
- 3. Can you please provide us with a copy of or reference to any key literature that shows the following relationships between any of the following (Fill in a Type of Paper cover sheet for each paper and/or reference)**
 - a. Different types of curriculum on: pedagogy learning outcomes social outcomes assessment
 - b. Different types of assessment on: pedagogy learning outcomes social outcomes curriculum
 - c. Different types of pedagogy on: learning outcomes social outcomes
 - d. Different types of learning outcomes on social outcomes
- 4. What do you consider to be the key trends and issues arising from literature with which you are familiar to inform policy (attach another sheet if required)**

Please attach a 'Type of Paper' cover sheet (separate page) to each paper you submit or each reference provided

Type of Paper cover sheet⁴

Author(s)

Title

Publisher

1 This paper is (please tick)

- a informed by empirical studies
- b informed by case studies
- c informed mainly by theoretical argument
- d a major review paper

2 This paper informs (please tick)

- a. Different types of curriculum on pedagogy learning outcomes
social outcomes assessment
- b. Different types of assessment on pedagogy learning outcomes
social outcomes curriculum
- c. Different types of pedagogy on learning outcomes social outcomes
- d. Different types of learning outcomes on social outcomes³ **Key issues and trends in the**

paper (please fill out)

⁴ Copy and fill in a form for each paper/reference provided.