MONASH UNIVERSITY

THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ON..... 26 July 2005.....

.....

Sec. Research Graduate School Committee

Notions of Work-related Skills and General Abilities:

The Generic Skills Debate

and the

Whole-school Assessment of Generic Skills

Doug McCurry

B.A. Hons. Dip. Ed.

Monash University Faculty of Education

Date submitted: 7.7.2004

Abstract

Notions of Work-related Skills and General Abilities: The Generic Skills Debate and the Whole-school Assessment of Generic Skills

Doug McCurry

In this thesis I sketch the development in the early 1990s of the Mayer Key Competencies in Australia and I compare them with other notions of work-related skills developed internationally at about the same time.

I focus on the ambitious assessment proposals of the Mayer Committee and the different views and arguments that were made about those proposals. Having reviewed the Key Competencies debate, I analyse notions of competence and ability and distinguish the generic Mayer competencies from other ideas of competence and competencies. As a result of this analysis I argue that the Mayer Key Competencies must be seen as generic abilities rather than specific competencies. I also argue that the Key Competencies are not achievements but rather that they are best thought of as aptitudes that predict abilities to learn new things.

I consider 100 years of research in psychometrics and cognitive psychology about generic abilities with particular attention to the work of H. Gardner, R. J. Sternberg, J. B. Carroll and S. J. Ceci. As a result of this analysis I sketch a model of cognitive abilities within an overarching model of performance.

I then turn from the theoretical and research literature on generic abilities to consider what happened to the assessment proposals that were referred by the Mayer Committee to the curriculum and assessment authorities in the various states and territories of Australia. I undertook reviews of Key Competencies assessment issues for the Commonwealth and the Victorian governments in 1996, and in that work I proposed a regime for school-based assessment of levels of performance on the Key Competencies. With Commonwealth Government support this proposal was trialled in 10 secondary schools with 110 teachers assessing 350 Year 11 students in 1997. The aim of the trial was to have students separately assessed by groups of teachers to see what degree of agreement there was between different teachers from different subject areas.

Through this trial I developed the notion of whole-school assessment in which all the teachers of a student contribute to a single, integrated report on the generic abilities of a

student. In the trial I conceptualised and implemented a cost-effective method of producing a collective view of a student from all the teachers of that student.

The assessment trial showed that teachers could in most cases make global, impression judgements of Key Competencies performances with no more that three minutes formal reflection per student. The judgements made by different teachers were quite consistent with each other, and as a result they can be validly and reliably used to develop an overall report for a student. The teachers participating in the trial judged the assessment procedures to be efficient and cost-effective.

Analysis of the assessments of teachers participating in the trial shows that pairs of teachers (from any combination of subject areas) typically produce an acceptable level of agreement in about 90% of cases. The further work I have done on the whole-school assessment of Key Competences has demonstrated on three other occasions that the assessment process I have developed is practical and useful.

Contents

Table of Figures

Table of Tables

Abbreviations

ACACA	Australasian Curriculum Assessment and Certification Agencies
ACER	The Australian Council for Educational Research
AEC	Australian Education Council
ANTA	The Australian National Training Authority
CBT&A	Competency based training and assessment
CCC	Cross-curricular Competencies Project of the OECD
CCE	Common Curriculum Elements
DEETYA	The Commonwealth Department of Education, Employment, Training and Youth Affairs
DeSeCo	Definition and Selection of Competencies project of the OECD
DETYA	Department of Education, Training and Youth Affairs
DomProAt	Domain and Process-related Differentiation of Cognitive Ability Nested in Conation and Affect
DomProC	Domain and Process-related Model of Cognition
DomProPD	The DomProPD Model of Cognition for Particular Domains
IAEP	International Assessment of Educational Progress project
IEA	The International Association for the Evaluation of Educational Achievement
INES	International Indicators of Educational Systems
KCs	The Mayer Key Competencies
MCEETYA	Ministerial Committee on Education, Employment, Training and Youth Affairs
MOVEET	Ministers of Vocational Education, Employment and Training
NCVER	National Centre for Vocational Education Research
NOOSR	The National Office of Overseas Skills Recognition
NVQs	National Vocational Qualifications
OECD	The Organisation of Economic Cooperation and Development
QCA	Qualifications and Curriculum Authority
QERC	Quality Education Review Committee
SBCKLAP	School based Key Competencies Levels Assessment Project
SCANS	Secretary's Commission on Achieving Necessary Skills
SCU	Socio-cultural understanding
STU	Scientific and technological understanding
VBOS	The Victorian Board of Studies
VET	Vocational Education and Training
VCAA	Victorian Curriculum and Assessment Authority
WSA	Whole-school Assessment

STATEMENT ABOUT PROVENANCE

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other institution.

The thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Acknowledgements

The research and development work described below was initiated within a number projects funded by the Commonwealth Department of Education Training and Youth Affairs and the Victorian Board of Studies (and their subsequent incarnations). The first of these projects was the review of the assessment activities undertaken in the Key Competencies Pilot Phase (McCurry 1996). In this project I participated in the Australian Curriculum, Assessment and Certification Authorities' project on the assessment of the Mayer Key Competencies (Edwards and Smith 1996), and then I undertook a review of the possibilities for assessment of the Key Competencies in Victoria for the Board of Studies (McCurry 1996). Subsequently I was involved in the National Industry Education Forum's Portfolio Project (MCEETYA 1996). I designed and directed the School Based Key Competencies Levels Assessment Project invited by and funded by the Commonwealth Department of Education, Employment and Training that generated the empirical data reported here (McCurry and Bryce 1997). I also undertook a review of the place of the Key Competencies in the Curriculum and Standards Framework, the courses in the Victorian Certificate of Education and the Vocational Education and Training courses within the Victorian Certificate of Education for the Victorian Board of Studies (McCurry 2001). In that project I analysed the options for Key Competencies assessment and recommended a trial of the model of whole-school generic skills assessment that was developed through the empirical aspect of this research (McCurry and Bryce 2002). This field trial took place in some 20 Victorian schools in 2002, 40 schools in 2003 and 50 schools in 2004.

The support of the Commonwealth Department of Education, Employment and Training and Youth Affairs (DEETYA) and the Victorian Board of Studies (VBOS) for this work is gratefully acknowledged.

The fruitful collaboration I have had with Dr. Jennifer Bryce of the Australian Council for Educational Research (ACER) in the execution of the School Based Key Competencies Levels Assessment Project (SBCKLAP) and the Victorian Board of Studies Key Competencies consultancy of 2000-4 is gratefully acknowledged. The original data analysis by Mr Greg Macaskill of ACER of the SBKCLAP data and his work on the Whole-school Assessment computer application is gratefully acknowledged.

An Overview

The research and development work reported here is conceptual and theoretical on one hand, and experimental and empirical on the other. In this thesis I sketch the development of the Mayer Key Competencies[1] in Australia, and compare them with other notions of work-related skills developed internationally. I analyse notions of competence and ability and the different theories of ability that have developed in psychometrics and cognitive psychology in the C20. I trace the debate about the proposals for the assessment of Key Competencies and the attempts to assess them within the Commonwealth funded program called The Key Competencies Pilot Phase (MCEETYA 1996). Finally I outline a notion of whole-school assessment of Key Competencies and report the results of a trial of a method for undertaking such whole-school assessment of generic skills and attitudes.

The last part of this thesis is concerned with the assessment of certain kinds of abilities, or notions of ability, using a certain kind of assessment and a specific methodology. It is concerned with notions of generic or cross-curricular abilities and with collaborative or whole-school judgements of such abilities. There is a specifically empirical basis to my work on the Key Competencies (it is a certain kind of assessment in certain contexts) and it was undertaken at a certain point of time in the history of Australian education. Something of this educational and historical context will be adverted to in outlining the specific empirical research, but that very interesting story (Where did the Key Competencies come from? Why? What happened to them? Why?) will only be glanced at in passing. The centre of this work is the theoretical or technical arguments that grew up around the Mayer Key Competencies rather than the educational and political arguments, important and interesting as those educational and political arguments are.

However this thesis is not, to me at least, merely theoretical and technical. It has been, and still is, practical development work. Through this work I have been concerned to design and test a realistic and cost-effective mode of assessing generic skills, and in the process I have found myself conceptualising the notion of generic skills assessment as what is called here 'whole-school assessment'. The notion of whole-school assessment is implicit in the Mayer Committee report, but it was inchoate in that work. I have tried to crystallise this implicit concept and then find a practical and cost-effective way of undertaking such school-based assessment.

[1] For convenience I capitalise the term key competencies when referring to the Mayer Key Competencies.

In working on the Key Competencies I have found myself in debates that touch on important matters of cognitive and assessment theory. The theoretical debates outlined below are part of my thinking about such theory, and reviewing this theory is an attempt to bring my own thinking about notions like the Key Competencies into a fruitful relationship with 100 years of thinking about abilities within psychometrics and cognitive psychology.

Since 1988 I have been a member of the Measurement Division of Australian Council for Educational Research (ACER), the major centre of the psychometric approach in Australia since the 1930s. In this thesis I have attempted to bring to bear what I have learned there about psychometrics and cognitive psychology on the challenge of whole-school assessment of generic skills. The specific topics dealt with in that development and implementation work must be approached from, or understood against, the backdrop of an understanding of human abilities and notions of assessment of intellectual abilities in general.

The psychometric approach to ability has been a matter of sometimes bitter and sometimes overtly political controversy. The various criticisms made of the psychometric approach will be outlined below. There is some real substance to these criticisms, but we will see there are methodological strengths and substantive understandings of ability developed within the theoretical and empirical aspects of the psychometric tradition. In my view, this body of work can and should contribute to developing notions of generic and work-related skills. It should also contribute more than it has to the way educators work with learners, but that is another story.

Although the outcome of my work on the Key Competencies is experimental or empirical, no apology is necessary for beginning with and framing the empirical work with the contextual and conceptual issues. As the following discussion suggests, too much assessment is done in an ad hoc, atheoretical or merely operational fashion. And the assumptions of much assessment are commonly less rigorously conceptualised than they might and should be.

1

The Development of the Mayer Key Competencies

The Economic Background to the Development of Notions of Generic and Work-related Skills

Perceptions of the economic circumstances of developed countries seem to have become less buoyant and more pessimistic over the last 30 years. After the optimism and boom times of the 1960s and early 1970s, there was the oil crisis and the rise of unemployment in the mid 1970s. The late 1970s were typified by stagflation. The 1980s saw recession and substantial levels of unemployment in developed countries, followed by increasing globalisation of trade and an intensification of international economic competition. A 'restructuring' of Australian industry seemed an imperative in the 1980s, and as this economic reorganisation took place, an increasing desire for change in the vocational and general education systems became evident in Australia and other similarly developed countries (Mayer 1992; Curtis and McKenzie 2001; Kearns 2001).

This research is concerned with notions of generic and work-related abilities that developed in Australian industry and education during the 1980s and 1990s. These developments in Australia (and the international parallels discussed below) arose from the economic situation sketched above. The major outcome of the thinking about generic and work-related abilities in Australia was the notion of the Key Competencies. These constructs were formulated under the sponsorship of the Commonwealth Government by a committee chaired by a prominent businessman, Eric Mayer. The work of the Mayer Committee (as it came to be known) was a central element of an attempt to develop nationally consistent approaches to education and training in Australia in the late 1980s and 1990s. A similar movement to national consistency can be seen in the development of nationally agreed goals for schooling and the collaborative development of national curriculum statements and profiles (Piper 1997).

The Report on the Outcomes of the Key Competencies Pilot Phase (MCEETYA 1996) described the development of notions of generic and work-related skills in Australia and overseas in the first half of the 1990s as a result of economic changes in the 1980s:

> Australian moves to examine the workplace relevance of school learning took place against the backdrop of a worldwide movement in the same direction, at least in most OECD countries. During the 1980s, profound changes in the economic circumstances of most industrialised societies, including accelerated technological change, and an accompanying shift in policy sentiment, led to a universal focus on the potential contribution of education to national well-being and, in particular, economic well-being. The emergence of higher levels of structural unemployment among young people also lent urgency to any reconsideration of training and education in the post-compulsory years. Existing systems of general education were re-examined to see whether or not they made an adequate contribution to national goals in a rapidly changing work environment (MCEETYA p.11).

This report retrospectively traced the beginning of these developments in the desire of the Commonwealth Government to develop a range of nationally consistent educational changes:

> In Australia, the Commonwealth Government responded to the emerging policy agenda with a series of developments aimed at strengthening the links between schools, industry and the community (Strengthening Australia's Schools, Dawkins, 1988), and at providing a range of training opportunities in the immediate post-school years and beyond (Industry Training in Australia: The Need for Change, Dawkins, 1988). At the same time, State, Territory and Commonwealth Ministers of Education were developing a set of common goals for Australian schools, and moving towards developing a program of interstate cooperation in educational reform, particularly curriculum reform (MCEETYA p.11-12).

Under the heading 'Putting General Education to Work' (Mayer 1992), the report of the Mayer Committee described the impetus behinds its recommendations as a pressure to change Australian workplaces 'in response to the need to improve productivity and compete with world's best practice in international markets' (Mayer p.viii). According to the Mayer Committee, these changes would become more rapid 'as the Australian economy becomes even more closely integrated with the economies of the region and our major trading partners' (Mayer p.vii). They described the changes in workplaces as a

> ... move away from specialised jobs and separate functions towards more broadly defined work roles and organisational structures that provide for devolved and shared responsibility for planning and decision making. Greater value is being placed on factors such as creativity, initiative, being entrepreneurial and being able to think critically about how to improve work practices (Mayer p.vii).

The Generic and Work-related Skills Movement

The economic restructuring of the 1980s required changes in work and work practices, and the need for such changes became the impetus for an examination of the relationship between education and work in many countries. The prime outcome of this examination might be described as the generic or work-related skills movement of the 1980s and 1990s that can be seen in Australia and other (particularly English speaking) countries. A set of Necessary Skills was identified by the federal government in the United States in 1990 (SCANS 1992). Core Skills were identified in Britain as part of the General National Vocational Qualification in 1990 (CBI 1989; SCAA 1996). New Zealand identified a set of Essential and Generic Skills for integration across the education system in 1990 (NZQA 1993). A set of Employability Skills was identified in Canada in 1992 (CBC 1992; CBC 2000). The Key Competencies were developed through the sponsorship of the Commonwealth Government in Australian in 1992. Generic, cross-curricular outcomes were identified for education in South Africa in 1995 (WCDE 1997). The Organisation of Economic Cooperation and Development undertook a project on the measurement of cross-curricular competences between 1992 and 1996 (OECD 1997). This project was followed by a project on the definition and selection of key competencies between 1996 and 2002 (OECD 1999).

The Momentum for Generic and Work-related Skills

While there are distinctive aspects to the genesis and nature of each of these different initiatives, in general they represent an increasing international interest in notions of generic and work-related skills. The impetus for the development of such skills was commonly attributed to the following economic and social factors:

- economic recession and unemployment;
- increased economic competitiveness and economic globalisation;
- changes in the nature and organisation of work (and dissatisfaction with the ability of new employees to adapt to the workplace or make the best use of technical skills);
- the speed of economic, technological and social change and uncertainty about the future;
- a desire to attain world's best practice in business and industry; and
- the growth of a culture of public accountability and concern about educational standards.

The generic and work-related skills movement developed from a desire to:

- improve the skills of workers;
- make the young more employable;
- reform general and vocational education; and
- make general and vocational education more appropriate to the workplace by making vocational education more general and general education more vocational.

The revitalisation or reform process envisaged by the generic and work-related skills movement was sponsored by governments with the prompting and significant support of both business and unions (ACTU 1993; Carmichael 1993; NIEF 1995; Stanton 1995). Such an alliance was very formidable indeed.

The straitened economic situation of the 1980s saw an increase in criticism of the education and training system in Australia. In some respects such criticism was not new or surprising, but there was a new edge to the questioning of the aims of education in the 1980s in Australia. There was a new emphasis on the need for the education systems to serve the economy and to better prepare young people for work. A number of particular pressures on education can be discerned in the more general social and economic imperatives described above. Rapid technological and social change, in particular, came as a tacit challenge to the goals and practices of education in Australia. The 1980s saw in Australia:

- increased dissatisfaction with the quality of school leavers and a desire for greater emphasis on 'work readiness' in schooling;
- increasing emphasis on basic skills, particularly the basic skills of literacy and numeracy;
- concerns about the adaptability and flexibility of young people;
- a certain dissatisfaction with the more academic or tertiary oriented curriculum of upper secondary schooling;
- increasing recognition of business and industry as an interested party in the nature and outcomes of education;
- questions about the aims of upper secondary schooling provoked by increased retention of students; and
- increased expenditure on education leading to increased demands for satisfactory returns on the investment in education.

In summary, it seems that uncertainties about the present and the future led the community to look more critically at the education system in the 1980s and 1990s. Not unreasonably, some educators felt that the education system was being held responsible for (even scape-goated for?) the broadest of economic and social difficulties. The increasing prominence given to work-related skills and the suggestion that education should give greater emphasis to such skills was viewed with real disquiet by some educators and educational administrators (Collins 1993).

The Impetus in Australia

Laurie Carmichael, a prominent union leader, was an important advocate of vocational training reform and the Key Competences in Australia. He described the workplace imperatives behind the reform of vocational training and the Key Competencies as arising from the pressure from global competition, a computer-based technological revolution, and the emergence of workgroups and teams (Carmichael 1993).

Anne Borthwick, the Secretary to the Mayer Committee, gave particular emphasis to the way the development of the Key Competencies was a response to changes in the workplace (Borthwick 1992):

> **How workplaces are changing**
>
> This focus emerged from consideration of the changes taking place in Australian workplaces to improve productivity in an effort to match, and improve on, world best practice. ... At the heart of these changes is the realisation that workplaces that have participative management styles, shared goals, multi-skilled workers and flat management structures ("new workplace cultures") are not only more productive than their converse ("old workplace cultures") but also cannot operate without employees whose personal and social skills are valued as highly as their technical competence.

While the impetus behind the development of the Key Competencies seems clear, there are some fundamental uncertainties, even confusions, about the notions of generic and work-related skills, and in particular about the notions of competence and competencies. As a result the Key Competencies have some potentially contradictory antecedents.

Perceptions of the Key Competencies were partly coloured by the rhetoric of competency-based training (CBT) that developed in the Vocational Education and Training sector (VET) during the 1980s. These perceptions were also fostered by a growing interest in the development of generic skills and a focus on outcomes in general education, as represented in the Quality Education Review Committee (QERC) discussed below (QERC 1985).

The Notion of Competence in Trades and Professions

The term 'competence' and the phrase 'competency-based training and assessment' (CBT&A) have a long, diverse and rather chequered history in educational discourse (Bowden and Masters 1993). The terms first became common in the 1960s in the United States as an expression of the movement away from norm-referenced to criterion-referenced assessment. The terms were also associated with the mastery learning approach to specifying and enumerating specific micro-skills and with the definition of minimum competence. This approach, at its most extreme, can lead to the specification of competence in terms of discrete performance objectives that could be ticked off a checklist. Some educators in the 1960s and 1970s saw such a competency checklist as an ideal.

The terms CBT&A have also been associated with a movement away from the inputs of education in terms of courses undertaken or time served in on-the-job training and apprenticeship. Instead of focussing on inputs, CBT&A was to be focussed on what a trainee had 'learned and could do'. This CBT&A was thus said to be performance or outcomes-based. It entailed judgements about the level of performance reached, and was particularly focussed on the threshold of competence. Such notions of CBT&A were a significant challenge to the traditional trade training system that had a time-served view of completing apprenticeships, and that aimed to reduce trade knowledge to discrete, demonstrable performance objectives.

There was general acceptance that CBT&A represented a valuable emphasis on quality in trade training, and the same kinds of views began to be applied by some to professional training and then to general education. The National Office of Overseas Skills Recognition (NOOSR) is a body funded by the Commonwealth Government with responsibility for assessing the qualifications and skills of those trained overseas who want professional registration in Australia. As well as being responsible for the assessment of trade and professional qualifications obtained overseas, in some cases NOOSR makes direct assessments of the skills of the applicants for recognition. In the early 1990s NOOSR began to promote the notion of competency-based assessment (CBA) as a way of dealing with the difficulty of assessing the levels of performance of overseas trained professionals.

According to NOOSR, CBA for the professions would entail a specification of the skills and knowledge of a profession and the specification of a minimum level of competence against which applicants would be assessed (Gonczi, Hager et al. 1990; Masters and McCurry 1990; Gonczi, Hager et al. 1992; Gonczi, Hager et al. 1993; McCurry and Bryce 2002). Such specifications of knowledge and skills and levels of competence did not exist for the

professions, and the suggestion that they should be developed was greeted with considerable scepticism and resistance by some professionals and educators of professionals (Wolf 1991; Barnett 1994).

The notion of competence defined by more or less specific performance objectives that had gained prominence in trade training was not readily applied to the complex and multi-faceted expertise of the professions. There were objections to the notion that CBT&A as micro-specification was realistic or appropriate for high-level professional expertise. Some professions and professionals argued that competence would have to be understood differently in the professions from the way it was understood in trade training and certification.

The Relational Model of Competence

The nature of the dissatisfaction with proposals for CBT&A for the professions can be seen in the review of Bowden and Masters (Bowden and Masters 1993) of the implications of CBT&A for higher education undertaken for the DEETYA and NOOSR. In their report the reviewers warned about the possible negative consequences of CBT&A:

> Indeed, if there is one continuing danger both to university education and to professional practice, it is the possibility of bureaucratically inspired external interference in the planning and conduct of professional education (p.ix).

In support of their view of 'competencies', Bowden and Masters presented what they called a relational model of observable practice and underlying capacities as a way of outlining the difficulties involved with generic competencies. This model, and the reasoning offered to support it, give a convenient introduction to some of the issues that are raised to challenge notions of generic skills and the Key Competencies.

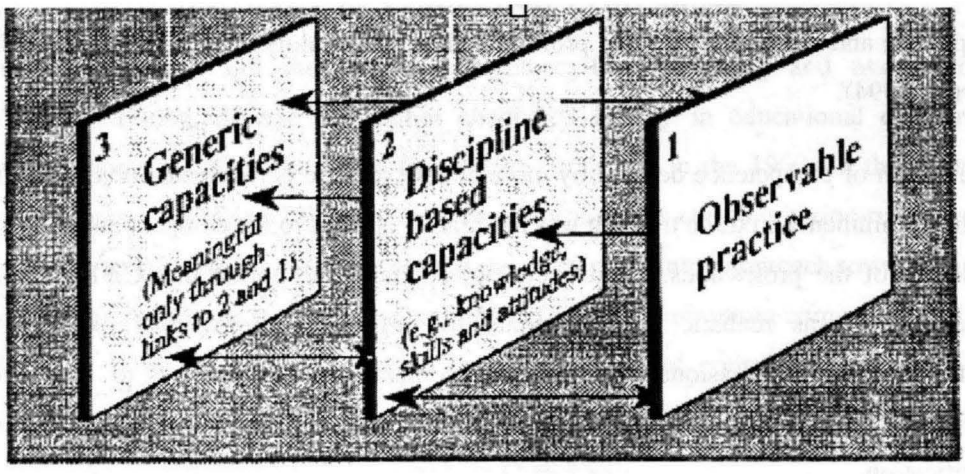


Figure 1 Bowden and Masters' Relational Model of Competence

This 'relational model' seeks to explain 'observable practice' in terms of two underlying capacities. These capacities are distinguished on the basis that one is 'discipline based' and the other is 'generic', but the nature of these different capacities is unclear. According to Bowden and Masters, the notion of generic capacities is only meaningful through links to discipline knowledge that then manifests itself in observable practice:

> Generic capacities derive their meaning through interaction with some concrete knowledge domain or real-world practice. Solving real-world engineering problems and solving difficulties with personal relationships in a social setting, for example, don't lend themselves to simple algorithms, let alone the same algorithm. The idea of teaching problem solving in an abstract, context-free way doesn't bear scrutiny; all learning is contextual (p. 157).

What is meant by a generic capacity having meaning is not clear in this statement or in the text as a whole, but the overriding significance given to 'concrete knowledge' is clear. The notion of general problem solving (the recurrent target for critics of generic skills) is presented by Bowden and Masters as wanting to apply the same 'simple algorithms' to any and all problems. The writers then assert that the very idea of teaching problem solving in an abstract and context-free way is absurd because all learning is contextual. A little introductory analysis of what will be recurrent issues in this discussion would be appropriate at this point.

In offering to consider the notion of competent performance, Bowden and Masters claim not only that knowledge is important and even crucial in explaining performance, they also claim that the notion of generic capacities is only meaningful through concrete and domain specific knowledge. It seems to Bowden and Masters that one might 'imagine' generic

capacities, but they equate the imagination of such constructs with the claim that the capacities would be taught in an abstract and context-free way:

> At a third level, a set of generic capacities can be imagined. These might include problem solving, critical reasoning, planning and organisation skills, and inter-personal and communication skills. Capacities of this kind can be stated as desirable outcomes independent of discipline or course of study. However, as argued earlier in this Chapter, these generic capacities are unlikely to be transferable across disciplines in the sense that a good 'problem solver' would be a good problem solver in any context. There is evidence that generic capacities are usually if not always highly context dependent. Nevertheless, it seems useful to separate out in this model the concept of generic capacities identified as common (or core) concerns across a spectrum of educational and training programs. However, they are only meaningfully assessable in terms of a specific body of knowledge or practice, as indicated by the double-headed arrows in the Figure (p. 159).

As we will see later, it is commonly claimed that there is a lack of transference of generic skills across contexts. Bowden and Masters make this claim, and they claim that generic capacities are usually context dependent. We are told that it seems useful to separate concepts of generic capacities as common concerns across educational and training programs, and the writers seem to see the generic capacities they are discussing as vaguely generalised goals ('desirable outcomes'). The writers then claim that generic capacities are only meaningfully assessed in a 'specific body of knowledge'.

What it means to say that generic capacities are only assessable in a specific body of knowledge is not clear nor is it explained, but some implications of this reasoning are clear. If generic capacities are only assessable in a specific body of knowledge then it would seem that such generic skills cannot be meaningfully separated from a specific body of knowledge. If generic capacities cannot be assessed separately from a body of knowledge then they are un-assessable as such.

By the time Bowden and Masters addressed these issues in 1993, the terms 'competence' and 'competencies' had became so contested and controversial that they eschewed the terms and referred instead to 'generic capacities'. The Mayer Committee designated a set of generic skills that were obviously different from the micro-skills approach to CBT&A as understood in trade training, but the use of the terms competence and competencies by the Mayer Committee for such generic skills enmeshed their work in controversy about the meaning and nature of competence and competencies in trade and professional training and certification. This ill-omened choice of a label (perhaps dictated by no more than the lure of alliteration?) has bedevilled much work on generic skills in Australia and internationally.

The qualifier 'Key' added to the term 'Competencies' was presumably supposed by the Mayer Committee to distinguish the macro level of the Mayer competencies from a micro level performance objectives approach, but the distinction was not always apparent or appreciated. By using the term competencies to describe their constructs the Mayer Committee was courting confusion of their generic skills with CBT&A. The Mayer Committee did not emphatically enough distinguish the generic nature of the Key Competencies from the micro skills view of competency specification, and, as a result, did not clearly enough distinguish such generic skills from simplistic notions of assessing what trainees 'know and can do' through discrete performance objectives. It will be argued below that generic skills or general abilities like the Mayer Key Competencies cannot be usefully understood as a matter of more of less specific skills (let alone a body of knowledge or 'understandings') of the kind enshrined in the deceptively simple formula about what people 'know and can do'. The Key Competencies of the Mayer Committee are generic skills or general abilities that are not, or cannot be, understood as more or less finite and specific knowledge and skills. They are very generalised notions that must be understood as trans-domainal. They cannot be understood as particular performances and defined in terms of particular tasks or performance objectives.

General abilities are not matters of particular knowledge and skill that can be discretely demonstrated. They are general attributes of individuals, or capacities attributed to individuals, and as general attributes, such capacities have a trans-domainal and hence predictive potential. If such abilities are transdomainal they can be applied in unfamiliar domains.

The notion of such generalised characteristics or attributes is, of course, not unproblematic or uncontroversial. The controversy provoked by such notions will be explored below, both in terms of the development of the Key Competencies and more generally in terms of various theories of cognition.

The Karmel Quality of Education Review

One of the early uses of the term 'competence' in Australian education was in the Quality Education Review Committee (QERC) chaired by the economist Peter Karmel that reported in 1985 to the Commonwealth Government. This review was clearly concerned with competencies as general abilities in general education rather than the specific skills of CBT&A.

The QERC review addressed the 'outcomes' of general education, and it described competence in the most generalised terms as 'the ability to use knowledge and skills

effectively to achieve a purpose'. For this review the notion of competence was practically concerned with nothing more specific than the outcomes of education. The QERC Committee identified the following very general 'competences':

- acquiring information;
- conveying information;
- applying logical processes;
- practical tasks; and
- group tasks.

By 1989 the notion of CDT&A had gained significant attention in vocational training, and the Commonwealth, State and Territory governments and the peak industrial bodies endorsed it in an agreement called the National Training Reform Agenda. This agenda formally introduced the concept of a 'competency-based approach' to vocational education and training to reflect the standards of education and training required for employment.

The distinction between competence as the specific skills and knowledge of a vocation and the generic abilities needed for a whole range of vocations (and for life in general) was often blurred, but the interest in 'competencies' as the generic outcomes of education and training continued to grow.

The Finn Report

In 1990, Commonwealth and State Ministers for Education and Training agreed that a national review of post compulsory education and training should be conducted. Brian Finn, the then Chief Executive of IBM Australia, was appointed as the independent Chair of the Review Committee. The Finn Committee, as it was known, produced a Report entitled 'Young People's Participation in Post-Compulsory Education and Training' in July 1991 (Finn 1992). Among other things, this report concluded that there are certain 'competencies' in a number of areas that all young people needed to learn in their preparation for employment. The recommendations of the Finn Committee were drawn with the broadest of brushes:

> Australia as a nation should be committed to providing for all its young people a program of education/training which prepares them for life as individuals, citizens and workers now, through the current decade and into the coming century (p.xv).

The Finn Committee recommended that:

> In recognition of Australia's need for a more highly skilled workforce which is able to operate more flexibly and with greater innovation at all levels, a framework of nationally agreed essential competencies and standards should be established and incorporated appropriately in all education and training programs for young people (p.xv).

The Finn Report listed six employment-related 'Key Areas of Competence' required to create closer relationships between education, training and work and as a means for industry to clarify its expectations of education, training and young employees. It called these things 'employment-related key competencies' and recommended that steps be taken to ensure all young people were able to develop these key competencies regardless of the education or training pathway they followed.

Table 1 The Key Areas of Competence of the Finn Report

Language and Communication Speaking Listening Reading Writing Accessing and using information	**Mathematics** Computation Measurement Understanding mathematical symbols
Scientific and Technological Understanding Scientific and technological concepts Impact of science and technology Scientific and technological skills	**Cultural Understanding** Australia's context Global issues World of work
Problem Solving Analysis Critical thinking Decision making Creative thinking Skills transfer to new contexts	**Personal and Interpersonal** Personal management Negotiating, team skills Initiative, leadership Adaptability to change Self-esteem Ethics

These areas of competence identified by the Finn Committee are general abilities that evidently contrast with the specific skill of a particular job or vocation.

The Carmichael Report

The recommendations of the Finn report were further developed in a review of vocational education chaired by Laurie Carmichael that reported in 1992. The Carmichael Report suggested that key competencies could provide a link between general education and vocational education and training through changes to industry certification. Some of these ideas were extended in a reform of the apprenticeship system called the Modern Australian Apprenticeship and Traineeship System (Carmichael 1992).

The Mayer Report

The Process

A Committee was established in September 1991 chaired by Eric Mayer (the former Chief Executive Officer of National Mutual Insurance) to develop the work begun by the Finn Committee on key competencies. The task of the Mayer Committee was to identify the key competencies and to develop a means of describing them that would provide a common reference point for curriculum and teaching in both the school and training sectors. It was also envisaged that the key competencies would provide the basis for a nationally consistent approach to assessing and reporting achievement in such competencies.

The development of the Key Competencies was undertaken by a broadly representative committee through a process of deliberation and consultation. All States and Territory Governments were represented on the Mayer Committee, and there were representatives from BP Australia, the Australian Chamber of Commerce and Industry, the Tourism Industry and the unions. The committee also included representatives from schools, vocational education and universities.

In September 1992 the Mayer Committee submitted its final report on Key Competencies, which was given, the catchy subtitle 'Putting General Education to Work'. The Report proposed a set of seven generic Key Competencies that young people need for effective participation in emerging forms of work and work organisation, together with principles to provide for nationally consistent assessment and reporting of achievement of the Key Competencies. State and Federal Ministers for Education and Training recommended that further work continue to refine the Key Competencies and to test their feasibility as part of the wider reform of the education and training system.

The Outcomes

The Report of the Mayer committee stated that 'there are certain essential things that all young people need to learn in their preparation for employment' (p.vii). It also stated that students should be able to develop these employment-related Key Competencies regardless of the pathway they follow through education and training in the post-compulsory years. The Report envisaged 'Records of Performance on Key Competencies' that would be issued to individuals as information for potential employers and for the purpose of credit transfer between different educational institutions. The Mayer Committee argued for pursuing national consistency in assessment and reporting through nationally agreed principles rather than common assessment instruments or methods. The report stated that such assessment could accommodate the diverse range of programs and settings of post compulsory education and vocational training.

The Response of Governments to the Mayer Proposals

Borthwick described the response of the government leaders of education to the Mayer proposals in the following terms (Borthwick 1992):

> At their meeting in Auckland, the Australian Education Council and Ministers of Vocational Education, Employment and Training adopted a number of resolutions in response to the Mayer Committee's report. The first of these underlined the place of the Key Competencies in reinforcing the role of general education as a foundation for vocational education and training.
>
> Ministers noted the importance of preparing young Australians for employment. This has always been part of the role of general education and is the primary purpose of entry-level training. The changes currently occurring in Australian industry to enable Australia to compete in international markets depend on developing a workforce capable of participating effectively in new forms or work and work organisation. This requires a renewed emphasis on the role of general education in providing the foundation for a multi-skilled, flexible and adaptable workforce and a greater emphasis on broad employment-related competencies in vocational education and training. Ministers welcomed the Mayer Committee report on Key Competencies as a significant contribution to addressing these education and training issues vital to Australia's future (P.6).

At the December 1993 meeting of State, Territory and Commonwealth education and training Ministers it was agreed that the States and Territories would collect information on the individual and collaborative approaches taken by States and Territories to the Key Competency proposals and report back to the April 1994 meeting of Ministers. This agreement seemed to represent an acceptance that the Key Competencies provided a nationally recognised framework for education and industry across Australia. But this general endorsement of the Mayer Committee recommendations did not involve, as Piper

noted, any recommendations concerning implementation, assessment or certification (Piper 1997).

An industry validation project was established to determine if the proposed Key Competencies adequately reflected industry and workplace needs. This project concluded that all seven Key Competencies are relevant to the existing and future needs of industry although further developmental work was needed.

Key Competencies Pilot Projects

The Commonwealth Government supported the Mayer recommendations by funding the States, Territories and peak bodies to conduct pilot work on Key Competencies. In 1993, the Commonwealth announced that it would provide 20 million dollars over three years to develop, trial and evaluate Key Competencies in general and vocational education systems. The objectives of this Key Competencies Program were to:

- enhance educational outcomes for all young people;
- promote the skill development necessary to enhance Australia's overall educational and economic competitiveness; and
- support the convergence of general and vocational education.

Some 70 different projects were funded under the program of 'pilot testing' of the Key Competencies in all States and Territories in 1994 and 1995. Some of the projects that underpinned the work reported here were funded under this program.

The Key Competencies

The Mayer Committee Report presented the Key Competencies as part of a worldwide interest in and research into the place of generic skills in education and training. The Committee argued that the Key Competencies are 'what people both know and can do' and they are 'essential for lifelong learning and effective participation in current and future society, education and training and work (paid and unpaid) and work organisation' (p.2). The Key Competencies (Table 2) were said to underpin success in employment, further learning and life in general.

These Key Competencies were defined by the Mayer Committee in the following way:

> ... there are certain essential things that all young people need to learn in their preparation for employment and that they should be able to develop these employment-related Key Competencies regardless of the pathway they follow through education and training in the post-compulsory years (p.2).

Table 2 The Mayer Key Competencies

The Committee has concluded that there are seven Key Competencies that all young people need to enable them to participate effectively in the emerging forms of work and work organisation.

KC1 Collecting, Analysing and Organising Information

The capacity to locate information, sift and sort information in order to select what is required and present it in a useful way, and evaluate both the information itself and the sources and methods used to obtain it.

KC2 Communicating Ideas and Information

The capacity to communicate effectively with others using the range of spoken, written, graphic and other non-verbal means of expression.

KC3 Planning and Organising Activities

The capacity to plan and organise one's own work activities including making good use of time and resources, sorting out priorities and monitoring one's own performance.

KC4 Working with Others and in Teams

The capacity to interact effectively with other people both on a one-to-one basis and in groups, including as a member of a team to achieve a shared goal.

KC5 Using Mathematical Ideas and Techniques

The capacity to use mathematical ideas such as number and space and techniques such as approximation and estimation for practical purposes.

KC6 Solving Problems

The capacity to apply problem strategies in purposeful ways, both in situations where the problem and the desired solution are clearly evident and in situations requiring critical thinking and a creative approach to achieve an outcome.

KC7 Using Technology

The capacity to apply technology, combining the physical and sensory skills needed to operate equipment with the understanding of scientific and technological principles needed to explore and adapt systems.

Although it used the term competence, the Mayer Committee was concerned to present what it called a 'broad definition' of competence and the Key Competencies, as distinct from specific knowledge and skills:

> Key Competencies are competencies essential for effective participation in the emerging patterns of work and work organisation. They focus on the capacity to apply knowledge and skills in an integrated way in work situations. Key Competencies are generic in that they apply to work generally rather than being specific to work in particular occupations or industries. This characteristic means that the Key Competencies are not only essential for effective participation in work but are also essential for effective participation in further education and in adult life more generally (p.ix).

The Mayer Committee set out to answer a daunting question: What is essential for effective participation in emerging work and work organization? In doing so it made a remarkable set of assumptions and claims, and produced an ambitious set of proposals.

The Key Competencies Assessment and Levels of Performance

The most ambitious feature of the Mayer Committee proposals was the emphasis given to the Key Competencies as the basis for assessment and reporting. The Committee also proposed that such assessment should be nationally consistent and that it be aggregated nationally. Not surprisingly, these ambitious assessment and reporting proposals came to be the most problematic and controversial elements of the Mayer Committee's work.

The Committee defined three levels of performance that were to apply to each competency:

> **Performance Level 1** describes the competence needed to undertake activities efficiently and with sufficient self-management to meet the explicit requirements of the activity and to make judgments about quality of outcome against established criteria.
>
> **Performance Level 2** describes the competence needed to manage activities requiring the selection, application and integration of a number of elements, and to select from established criteria to judge quality of process and outcome.
>
> **Performance Level 3** describes the competence needed to evaluate and reshape processes, to establish and use principles in order to determine appropriate ways of approaching activities, and to establish criteria for judging quality of process and outcome.

The assessment proposals of the Mayer Committee are fairly described as very ambitious to the point of audacity.

They proposed the:

- inclusion and assessment of generic skills constructs in all education and training programs;
- inclusion of such new and different curriculum and assessment constructs as Planning and organising and Working with others and in teams;
- opportunity for all students and trainees to be assessed on generic constructs; and
- aggregation of the results of the assessment for national standards monitoring.

There seems nothing to compare with the ambitiousness of these proposals in the history of Australian education. Even partial realisation of these goals would be arguably the single most substantial change ever made in Australian education. Such proposals make the notion of a national curriculum developed through national subject statements and profiles seem a very modest ambition. As was shown subsequently, both the national subject statements and profiles and the Key Competencies proposals have proved to be overly ambitious (Piper 1997).

The Key Competencies were promoted by the Commonwealth government with a certain amount of public relations fanfare, and substantial funding was used to 'seed' the initiative through the pilot program, but the audacity and momentousness of the Mayer proposals was hardly ever recognised or, at least publicly, admitted.

The Responses of Education Systems and Educators to the Mayer Proposals

The Key Competencies were and still are a significant challenge to the conventional approach and the status quo of education and training. An extreme or pessimistic response (and such were commonly vocalised but rarely recorded) would see the Key Competencies as undermining the conventional nature and structure of general education, and, on the other hand, undermining the job focussed specificity of vocational education.

Leaving aside the possible impact of the Mayer proposals on the current and historical practice in education and training, any implementation of the proposals would involve remarkable and unprecedented difficulties and challenges. A full realisation of the whole Mayer Committee vision should have seemed, and does in retrospect seem, a matter of remarkable fantasy. Or perhaps it was a kind of audacious bluff? What had really shaped

what happened can only be a matter of surmise, and no space can be given to such matters here. Taking the proposals at face value, one can only wonder about the forces that allowed such a fantasy to be entertained or that allowed the emperor's new clothes to be viewed without comment.

The Mayer Committee's call for education and training to give increased emphasis to preparation for work was backed by the most powerful sections of the community (ACTU 1993; DEETYA 1994; DEETYA 1994). These groups seemed to envisage a *de facto* national curriculum and an actual national assessment. The remarkably radical nature of the Mayer Committee proposal received an equally remarkable degree of initial acceptance from the state authorities at Ministerial level.

The Commonwealth Labor Government had sponsored the development of the Key Competencies but was not able to implement the proposed system. Implementation would be a matter for the state governments that had the constitutional responsibility for education. However the financial power of the Commonwealth could be used to encourage the adoption of the Key Competencies, and the very substantial sum of $20 million was made available for a project-based program that was initially described as implementation but was eventually described as a pilot program. Education Ministries, The Australian Curriculum Assessment and Certification Agencies (ACACA) and a whole range of organisations were willing to take (with very few strings attached) the funds offered by the Commonwealth, but the acceptance of the Key Competencies by the policy determining and implementing agencies as they took the Commonwealth funding is open to question. It seems as though the Commonwealth government believed that with some initial funding, the Key Competencies would take off like wild fire.

While the reactions of practitioners in classrooms to such proposals might seem to be the most fundamental of issues in the spread of the Key Competencies, the reactions of the educational bureaucracies comes to be the real determining issue in such top-down policy making. How would the educational bureaucracies, in particular those responsible for assessment and certification, respond to and deal with the challenge of the Key Competencies?

The educational bureaucracies could not reject or easily ignore such calls for change or innovation coming from powerful interest groups and supported by the heads of the Commonwealth and State governments, but there is reason for thinking the state educational bureaucracies that would be responsible for any implementation of the Mayer vision were at least cautious and uneasy about the Mayer proposals (Edwards and Smith 1996). There is

also reason to believe that the ACACA agencies were determined to be key 'stake holders' if there was to be any game (ACACA 1995). Even so it seems that the educational bureaucracies were actually suspicious and guardedly uncooperative in response to the Mayer proposals.

Educational bureaucrats were very cautious about what views they placed on record about matters apparently endorsed by their political masters. As a result, it is difficult to pin down the reactions of the educational bureaucracies to the Mayer proposals.

Assessment and certification turned out to be the most problematic and sensitive issues. The Mayer Committee report expressly sought advice from ACACA 'on the relationship of the Key Competencies and senior secondary school certification' (Edwards and Smith 1996). ACACA had to respond to the Mayer proposals, and the nature of these responses will be traced in Chapter 8 as they manifested themselves in the Commonwealth funded ACACA Key Competencies Project on Assessment and Reporting. This project shows some of the suspicion of and resistance to the assessment aspects of the Mayer proposals by the educational bureaucracies.

The most direct criticism of the Mayer proposals came from educational theorists and academics, and these criticisms were also reflected in the responses of ACACA. The criticisms of educational theorists and the reflection of these criticisms in the positions taken by ACACA agencies will be considered below. The response of ACACA to the Mayer proposals is a complex mosaic, but the reactions of these assessment bureaucracies can be summarised as openly cautious and non-committal but privately suspicious and hostile.

The political and educational tenor of the ACACA response to the Mayer proposals is a complicated story, and this discussion does not intend to deal with those wider aspects of the Mayer vision and the ACACA responses to it. Rather it aims to consider the more specifically theoretical and technical assessment challenges of the Mayer proposals and how the ACACA responded to them.

One of the extraordinary characteristics of the muted debate about the Key Competencies is that it was conceptual and philosophic rather than political or pragmatic. While there was little overt criticism of the Key Competencies from educational authorities, some of those outside the power structures were not constrained, and there was a good deal of criticism from academics. The disquiet of these educators was evident in more theoretical questioning and debate than in challenges to the social and political tenor of the Mayer proposals.

Running through various discussions of the Key Competencies is a consistent questioning of them conceptually and on the basis of 'theory' and 'research'. This theoretical and research-based questioning will be examined below. It will be shown that this questioning, and the pragmatic challenges of responding to the Mayer Committee proposals led to and shaped the empirical research reported in the last part of this thesis.

This review of the history of development of the Key Competencies shows that

- there was a tension between the tertiary and secondary education, government and business about the means and ends of education and training;
- the development of the Key Competencies was typified by loose terminology, lack of conceptual rigor and an atheoretical approach;
- the Key Competencies were challenged as inconsistent with what is known about learning and the mind; and
- there is a political tincture to the arguments of theoreticians and academics against the Key Competencies.

The discussion in the next section aims to compare the Key Competencies with other generic and work-related notions developed in other countries. The discussion will then turn to the conceptual and theoretical challenges offered in various quarters to the Key Competencies in Australia, and, in so far as they explicitly surfaced, from ACACA's consideration of assessment issues associated with them.

2

Notions of Work-related and Cross-curricular Abilities

The Nature of the Key Competencies

Competence as Specific Outcomes

'Competence' and 'competencies' are vexed terms onto which all manner of assumptions and presumptions have been projected.

Alison Wolf (Wolf 1995), one of the most vocal critics of the notion of generic competencies, offers the following definition of the term which she borrowed from American sources:

> Competence-based assessment is a form of assessment that is derived from the specification of a set of outcomes that so clearly states both the outcomes - general and specific - that assessors, students and interested third parties can all make reasonably objective judgments with respect to student achievement or nonachievenient of these outcomes; and that certifies student progress on the basis of demonstrated achievement of these outcomes. Assessments are not tied to time served in formal educational settings. (p.1)

Wolf uses this definition because it 'encapsulates most of the key features of competence-based assessment as it is actually being developed and promoted at present: most especially in vocational, technical and professional education and training in the UK' (p.1). The key features of this definition are the specificity of outcomes that define a competent level of performance. While the Key Competencies have a rhetorical allegiance to outcomes and levels of performance, they are evidently not a specific set of outcomes on which 'objective judgements' of achievement can be made.

The Meaning of Competence for the Mayer Committee

The Mayer Committee offered the following definition of 'competence':

> The term competence focuses attention on outcomes. It is about what people can do. The Committee adopted a broad definition of competence, which recognises that performance is underpinned not only by skill but also by knowledge and understanding, and that competence involves both the ability to perform in a given context and the capacity to transfer knowledge to new tasks and situations. This definition is consistent with the definition adopted by the National Training Board. (p.ix)

This broad definition of competence contrasts with the 'set of outcomes' described by Wolf. For the Mayer Committee the notion of competence means a focus on 'outcomes' and what people can do, but the 'broad definition' is not specific to work in particular occupations or industries. If we take a 'focus on outcomes' to mean no more than a concern with assessment of levels of performance, this is an intelligible claim. Certainly the Key Competencies are not occupation or industry specific, they give significant emphasis to assessment, and they could not be any much more generalised or generic without ending up with the notion of a single general ability.

The committee report claims that the Key Competencies 'focus on the capacity to apply knowledge and skills in an integrated way in work situations' (p.ix). However, what is meant by saying that they are 'integrated' is unclear. It is intelligible to say that the Key Competencies are integrated in that they are generic rather than an atomised checklist, but the definition claims that the Key Competencies are an integrated application of knowledge and skill. This claim might amount to no more than the desire to envisage that Key Competencies as kinds of activity. This desire seems to be the impulse behind all titles of the Key Competencies (apart from Cultural understanding[2]) beginning with participles. The emphasis given to knowledge and skill in the Mayer discussion should be viewed with care. One wonders what it means for the Mayer conceptualisation to see a generic skill as a matter of knowledge, and one wonders how the Mayer conceptualisation envisages a generic skill. We will return recurrently during this discussion to the question of the part knowledge plays in the Mayer conception and in the notion of generic skills.

There is one other aspect of the notion of the Key Competencies as an integration of knowledge and skills in application that is worth prefiguring at this point. To some interpretations it seems that the Key Competencies are integrated in application in such a way that they are inseparable from a particular application. This leads to a view that because they are integrated they cannot be specifically assessed. (We have seen an argument of this kind in the Bowden and Masters view of 'observable practice'.) Such an interpretation is clearly at odds with the Mayer vision, but we can see how the phrasing used by the Mayer committee in these matters spawns confusion and results in contradictions.

The final point of the Mayer definition of Key Competencies is that they are to be essential not only for work but also for education and for life as a whole. In one sense this is a clear enough claim. It might be expressed a little more precisely as the question: 'How might we envisage the core that is essential for competent participation in education, work and life?'

[2] Cultural understanding was considered as a Key Competency but it was not finally accepted.

This is, of course, a fundamental and significant question, but it may be one of those simple but overwhelming questions that are unanswerable. We should notice, for future reference, that this is the question addressed by what will in this discussion be called theories of ability. So in a sense the Mayer conceptualisation is seeking, without knowing it, to ask the basic questions about abilities and to construct a differential ability theory. As we will see below, the Mayer Committee did not recognise that they were constructing an ability theory; they did not consider theories of ability that have been developed; and they do not seem to have been aware of what such theories might show about core abilities and how they might best be understood. In other words, the Key Competencies were constructed in a theoretical vacuum.

The Mayer conceptualisation envisages collecting, analysing, communicating, planning, organising, working, using and solving as kinds or classes of activity. This emphasis on activity is familiar as a common desideratum in assessment rhetoric. Such rhetoric seeks to make notions concrete in that it views abilities as observable performances or behaviours. This impulse, and the accompanying rhetoric, have a significant history that will be glanced at in Chapter 6, but it is worth noting here that in its definition of a Key Competency the Mayer Committee had committed itself (unknowingly) to the approach to conceptualising ability that was pre-eminent in the first 50 years of the last century. This approach was a matter of dispute for the second 50 years of the last century (Gardner 1987; Hunt 1993).

To return to the definition, we are told that the Key Competencies are generic rather than specific, and they involve skill that is underpinned by knowledge and understanding. They involve the application of knowledge and skill in an integrated way and the capacity to transfer knowledge to new tasks and situations.

There is a good deal that could be unpacked and critiqued in these claims. The fundamental notion is that of a generic skill, that is a skill that can be applied in a range of contexts or domains, or even, perhaps, all contexts and domains. While the Committee does not foreground the notion of a generic skill, the definition of a Key Competency shows an evident awareness of the major challenges that have been made to claims about such generic skills, and the definition seems to be framed in response to various claims made against the notion of generic skills. The definition recognises the importance of knowledge in competent performance in response to claims that competent performance is based on the specific knowledge and skills of a particular domain. The definition also recognises the challenge that the knowledge and skills of a particular domain cannot be readily transferred from one domain or context to another (and that skills are not usefully thought of as trans-domainal).

Thus the Mayer Committee defines a Key Competency in terms of the capacity to transfer knowledge to new tasks and situations.

Although it seems that these definitions are framed in response to criticisms of notions of generic skills, the Committee offers no more response to these criticisms than to build the challenges into the definition. A notion is a Key Competency because it involves the application and transfer of knowledge and skills and is general. According to the Mayer Committee, there is no problem of foundation knowledge or with the transference of knowledge because the Key Competencies are defined as including foundation knowledge that can be transferred.

The Mayer definitions of competence and of Key Competencies attempt to either blandly pre-empt or fend off criticisms about the importance of knowledge in competence and the difficulty of transferring knowledge and skill from one context to another by entertaining an inadequate model of generic skills and of performance. The definition suggests a model of performance as a matter of learning something in one context and applying it in another. The Mayer conceptualisation cannot deal with the place of knowledge in competence and the difficulty of transferring knowledge across contexts because it simply assumes a model of learning and performance (and hence of generic skills) as the transference of specific knowledge and skills from one context to another.

While the Committee was evidently aware of challenges to the notion of generic skills, it does not actually deal with these challenges. We will consider these criticisms of the notion of generic skills in detail in later chapters, but we should note here that the Mayer Committee did not (and perhaps could not) adequately deal with such challenges. The Committee would have done its work more effectively if it had faced up to the challenge of foundation knowledge and transference of knowledge because it would have had to work out more rigorously and coherently what it might realistically aim to do.

Because the Key Competencies were not rigorously or coherently conceptualised, 'implementation' of the notions was plagued by confusion. It will be argued below that the Key Competencies can be conceptualised with reasonable rigor and that the challenges on the basis of foundation knowledge and transference can be rebutted. The limitations of the work of the Mayer Committee is partly a result of the membership of the committee itself, and partly a result of the process that was used to develop the conceptions.

The Mayer Committee Process

The development of the Key Competencies was a process of committee discussion and public consultation. The committee asked: 'What is required for work now and work in the future?' In answering this question they considered industry competency standards, similar work overseas, and the responses offered in two rounds of 'extensive consultation'.

The definitions of Key Competencies the Mayer committee offered are a matter of principle and desiderata shaped by the societal pressures described in Chapter 1. In offering its stipulations for Key Competencies, the Mayer Committee took some very definite positions about abilities, but it did so without elaboration or justification. The Key Competencies are a differential ability theory but they are not adequately conceptualised as such a theory. The committee assumes that there is something that can be usefully called, or thought of, as a competency that involves the integration and application of knowledge and skills, and that is generic to schooling, work and life in general. Various criticisms of these assumptions will be considered below.

The Key Competencies were seen to meet a need to give recognition to individuals for skills that are not recognised in most assessments. The committee also implied that the Key Competencies were an important means of monitoring educational standards and for increasing the accountability of education.

The Difference Between Finn Areas of Competence and the Key Competencies

There was a process of change, some might argue it was a process of refinement, between the Finn notion of Areas of Competence and the Mayer conception of Key Competencies. In some respects the Key Competencies were more conventional and conservative and in other respects more unconventional and radical than the Finn Areas of Competence. There is some reason for thinking that the educators had more influence in the Mayer Committee than in the Finn committee, or perhaps the practical and implementation focus of the Mayer Committee work made their deliberations more concrete and hard-headed. The Finn areas contrast with the Mayer Key Competencies in that the Finn Areas contain some broad domains of knowledge that are related to and can subsume different academic subject areas.

In reviewing the differences between The Finn Areas and the Mayer Competencies, Piper notes that the Finn Areas have a certain content orientation:

> It will be noted that while some of these identified areas are clearly process-oriented, and hence amenable to definition in terms of 'competence', others, such as scientific and technological understanding and cultural understanding, appear to be more content-oriented, although expressed in process terms (understanding) (p.50).

Piper describes the change from Finn to Mayer as a movement to a more process-oriented conception in the sense of being more activity or skill based than knowledge based:

> The Mayer Committee, charged with recommending on the implementation of the employment-related key competencies, redefined the Finn competencies and, in doing so, shifted their focus from a partly content-oriented formulation to an exclusively process-oriented formulation (p.50).

In particular, the Mayer Committee turned away from the domain-related constructs of Scientific and Technological Understanding and Cultural Understanding included in the Finn Areas of competence.

Perhaps the most evident difference between the Finn Areas and the Key Competencies is in what are commonly thought of as the more personal skills or attitudes. It is in the area of the Personal and Interpersonal that the Finn conception contrasts most clearly with academic disciplines and with the Key Competencies.

Personal and Interpersonal
Personal management
Negotiating, team skills
Initiative, leadership
Adaptability to change
Self-esteem
Ethics

The notions under the heading of 'Personal and Interpersonal' in the Finn conception are usually thought of as personal characteristics rather than cognitive skills, and as such they contrast with academic areas of study. The place of personal characteristics in the education systems of secular and pluralist societies, like Australia, is ambiguous because the goals of education commonly emphasise personal growth but the school curriculum is organised around more or less specific cognitive processing.

The goals of education commonly identify growth of character and ethical behaviour as key aims of the system, but the curriculum in secular and pluralist societies rarely involves matters of ethical and ideological instruction, and matters of ethics and ideology are certainly not matters that are dealt with in formal and certified assessment. In secular and pluralist educational systems, ethical and ideological matters are not at issue in formal assessment. Such assessments are intended to be assessments of cognition rather than matters of behaviour, character or ideology.

There is a marked discomfort with the notion of assessing anything other than cognition in secular and pluralist education systems. This discomfort can be seen in the curious discontinuity between the goals of our education systems and both the curriculum taught and the activities undertaken in schools. Something of this discomfort and equivocation about the ethical and ideological issues can be seen in the way the Mayer Committee deals with the personal and interpersonal areas of the Finn Committee.

The Key Competencies and Values and Attitudes

The conventional definition of vocational competence involves knowledge, skills and attitudes (Bowden and Masters 1993), but the Mayer conceptualisation defined competence as knowledge, skills and understandings. In defining the Key Competencies in this way the Mayer Report explicitly eschewed attending to values and attitudes. As justification for excluding the consideration of attitudes the Mayer Committee claimed that:

> Both the principles and characteristics the Committee has used to construct the set of Key Competencies preclude the inclusion of values and attitudes (p.13).

However, it is hard to see in what sense the principles and characteristics designated by the Committee preclude values and attitudes. The committee argues that:

> Personal qualities such as punctuality, initiative and honesty are a vital part of education and training and fundamental to work and community life. Many groups, particularly industry and parent groups, argued in consultations and submissions that the Key Competencies should encompass such values and attitudes. The Committee acknowledges and shares this commitment to the importance of values and attitudes. In doing so, however, it maintains the view that a set of Key Competencies can only contain those things which can be developed by education and training, which do not require some innate predisposition or adherence to a particular set of values and which are amenable to credible assessment. On these tests and, in some cases, the test of conceptual coherence, the Committee considers that attitudes and values fall outside the field of the Key Competencies (p.13).

These claims touch on a fundamental tenet of education in secular and pluralistic cultures. The committee claims that although attitudes are fundamental to education

they are not or cannot be developed by education and training and/or that they are necessarily innate. Both of these claims are at odds with the aims of most educational enterprises, and they are certainly at odds with National Goals of Schooling for Australia of 1999.

One can understand the intention of the Mayer Committee to exclude values and attitudes because they are ideological ('a particular set of values'), but one can only wonder what is meant by 'the test of conceptual coherence' undertaken by the Mayer Committee that excludes values and attitudes from their definition of competence. It is not clear in what sense the notion of initiative is less coherent than the notions of teamwork or planning and organising. Adding to the confusion, the Committee claims there is still an ethical dimension to a competency like Working with others and in teams, and in application of competencies in work places:

> There is, however, an ethical dimension to the Key Competencies and, given the way in which they are embedded in context, their development is likely to support the development of the attitudes seen as desirable. While work ethics such as 'a fair day's work for a fair day's pay' are not identified among the Key Competencies, other dimensions of personal and work ethics are evident. Working with Others and in Teams, for example, includes contributing to the good of the group or organisation and being ethical in one's dealings with others. In short, while attitudes are not competencies *per se* they are a function of particular work place settings which will be reflected in the development and application of Key Competencies (p. 13).

So on one hand the Mayer Committee explicitly eschews the assessment of values and attitudes and on the other hand claims that such matters are at least implicit in Working with others and in teams, Planning and organising activities and, perhaps, in Problem solving.

The Key Competencies and Outcomes-based Eduction

The Key Competencies offer themselves as outcomes rather than inputs, but they contrast markedly with the specific outcomes of CBT&A and the specific outcomes and indicators of the 'developmental continua' of the National Curriculum Profiles (Piper 1997). It is hard to see what being outcomes-oriented means in the Mayer conception other than the marked emphasis given to assessment. Further, it is hard to see how conceptions of generic skills can be usefully thought of as specific outcomes. By definition, generic skills are very broad generalisations about the capacity to perform. It will be argued below that competence is expertise, but that the Key Competencies cannot be thought of as expertise. The Key Competencies are aptitudes rather than achievements, and they cannot be readily envisaged as specific performances.

The Notion of Generic or Cross-curricular Capacities

The Mayer Report compares their Key Competencies with the Core Skills in the United Kingdom, Workplace Know-How in the United States and Essential Skills in New Zealand, saying that:

> In each case the initiatives have resulted in a set of statements about skills/know-how/competencies which are
>
> - not specific to any particular subject area, educational program, qualification or awarding body nor to any specific vocational task or career path, but which focus on generic attributes that can be learned through formal processes and that apply generally to working life;
> - common to both general education and vocational education and training;
> - concerned with outcomes, in each case defined as precisely as possible with various levels to indicate the variety of individual attainment (p. 13).

The frameworks developed by different work-related skills projects can be found in Table 3.

The Mayer Key Competencies grow out of, and are related to, an increased international emphasis on generic and cross-curricular capacities rather than curriculum-based learning. This international emphasis on generic and cross-curricular capacities involves:

- decreased emphasis on traditional subject areas;
- increased emphasis on the relationship of education and training to work;
- increased emphasis on the basic and generic skills of literacy and numeracy; and
- increased emphasis on the quality of outcomes from education.

The skills, competencies, attributes, characteristics, qualities or capacities sought and identified by this movement have been variously described as core, key, necessary; essential, generic; transferable, or employment or work-related. These matters have also been seen internationally, and in Australia, as related to that other major slogan of 1990s' education, 'lifelong learning'.

The emphasis on generic capacities elides or blurs subject areas and contrasts with calls for a core curriculum of essential learnings. Generic and work-related capacities are seen as relating theory and application or the academic and technological, and they are seen as relating education to work and life in general. In so far as they contrast with specific vocational skills, generic capacities show an interest in the claims for general education as distinct from subject areas or technical training.

Table 3 Various Work-related and Generic Skills Constructs

AUSTRALIA MAYER KEY COMPETENCIES	US SCANS WORKPLACE KNOW-HOW	UK CORE SKILLS	CRITICAL CROSS FIELD OUTCOMES SOUTH AFRICA	CANADA EMPLOYABILITY SKILLS PROFILE
Collecting, analysing and organising ideas and information Communicating ideas and information Planning and organising activities Working with others and in teams Using mathematical ideas and techniques Solving problems Using technology (Cultural understanding) **The Finn Areas of Competence** Language and Communication Mathematics Cultural Understanding Scientific and Technological Understanding Personal and Interpersonal Problem Solving	**3 foundation skills:** basic skills (reading etc) thinking skills personal qualities **5 competencies:** resources (time allocation etc) interpersonal information (acquires & evaluates info etc) systems (understands systems etc) technology **National Committee of Inquiry into Higher Education of UK** (The Dearing Report) The knowledge and understanding that a student will be expected to have upon completion of a program Key skills: communication, numeracy, the use of information technology and learning how to learn; Cognitive skills, such as an understanding of methodologies or ability in critical analysis; Subject specific skills, such as laboratory skills	communication personal skills numeracy information technology problem solving competence in modern language **QCA Key skills Britain** **Communication** Taking part in discussions and making presentations Reading and responding to written material Producing written material **Information technology** Preparing information Processing and presenting information Reviewing the use of information technology **Application of Number** Collecting and recording data Working with data Presenting findings **Working with Others** Planning activities Working towards identified targets **Improving Own Learning & Performance** Setting targets and action planning Following plan to meet target Problem Solving	Identify and solve problems in ways which display that responsible decisions using critical and creative thinking have been made. Work effectively with others as a member of a team, group, organisation, community. Organise and manage oneself and one's activities responsibly and effectively. Collect, analyse, organise and critically evaluate information. Communicate effectively using visual, mathematical and/or language skills in the modes of oral or written presentation. Use science and technology effectively and critically, showing responsibility towards the environment and others. Demonstrate an understanding of the world as a set of related systems by recognising that problem solving contexts do not exist in isolation	**Fundamental Skills** Communicate Manage information Use number Think & solve problems **Personal & management skills** Demonstrate positive attitudes & behaviours Be responsible Be adaptable Learn continuously Work safely **Teamwork skills** Work with others Participate in projects & tasks **OECD DeSeCo Key Competencies** Acting autonomously and reflectively Using tool interactively Joining and functioning in heterogeneous groups

The OECD Key Competencies Project

The most significant project in the area of generic and work-related skills was developed by the Organisation of Economic Cooperation and Development (OECD) entitled The Definition and Selection of Competencies project (DeSeCo) (OECD 1999).

The DeSeCo project grew out of the OECD program of 'measuring the output of educational processes' developed in the 1980s and early 1990s, but it contrasted markedly with the pragmatic and psychometrically based approach of other assessment projects undertaken by the OECD. Having analysed the previous OECD studies as the first step in the DeSeCo project, Salganik et al. concluded that previous competency-related large-scale studies within the OECD lack an 'explicit, overarching conceptual framework based on broad theories of what skills, knowledge, and competencies are and how they relate to each other' (p.5). DeSeCo was designed to work towards an overarching, conceptual framework through an 'interdisciplinary, international scientific approach', in close collaboration with ongoing OECD assessment programs, to:

- advance the theoretical underpinning of key competencies;
- provide a reference point for indicator development and interpretation of empirical results ;
- encourage an iterative process between theoretical and empirical work; and
- eventually provide feedback for education policy.

DeSeCo aimed to develop a common, overarching, theoretical framework for the identification of relevant skills and competencies (OECD 2001):

> This was done through an interdisciplinary scientific approach, within an international context, aimed at identifying those competencies needed for individuals to lead an overall successful and responsible life and for society to face the challenges of the present and the future (p. 1).

Although it began with a focus on theoretical underpinnings, the DeSeCo Program also aimed to provide 'a reference point for the development and validation of competence indicators and a basis for more accurate and appropriate interpretation of empirical results' (p. 1). The work of the program aimed to provide a resource for the process of defining and selecting key competencies, as well as for the development and understanding of future measures aimed at producing relevant indicators of skills and competencies. It was intended that the comparisons between the types of competencies required in various life situations and what people learn in different educational environments would provide feedback for education policy. It was also intended that relevant criteria for a successful evaluation of

educational systems could be selected on the basis of a theoretically founded framework. DeSeCo was based on the following positions:

- a broad definition of competence;
- key competencies exist and are particularly important;
- competence is more than the ordinary study programs of educational institutions; and
- there are multiple conceptual approaches to the competencies.

Although the emphasis given to theoretical and empirical research by DeSeCo markedly contrasts with the other work-related skills constructs described above, it recognised that 'the definition and selection of key competencies is still - and even perhaps first and foremost - a process of negotiation among various policy-makers, and not simply a question of scientific reflection' (p.4). The extent to which the work of DeSeCo ends up as a matter of policy and the extent to which it is a matter of science is a significant question.

DeSeCo set out to review a remarkably comprehensive set of considerations and questions. It addressed the most basic issues about human nature and society:

> What ideas about the nature of human beings and society should serve as a starting point for the identification of key competencies? What are the premises for a so-called successful life in various spheres of life from the perspective of both the individual and society? What are the underlying normative criteria for defining key competencies? Are there common denominators among the different viewpoints on this issue (Rychen et al. p.4).

DeSeCo considered how competencies might be designated and defined:

> What are the theoretical foundations, rationale, and selection processes behind the sets of key competencies? What are the political, social, and economic factors that influence the definition and selection processes of key competencies in different socio-economic and cultural environments, and how is this influence exercised? What is the role of scientific findings and scientific methodology in these processes?
>
> What are the convergences between the different processes and between the different sets of key competencies? Is there a significant amount of homogeneity across different environments, recognizing the heterogeneity of scientific methodologies and selection processes? To what extent does that heterogeneity need to be taken into account (Rychen et al p.4)?

As these statements demonstrate, DeSeCo took a fundamental approach to the issue of life-skills.

The DeSeCo Activities

After the initial undertaking of analysing previous competence-related work in the OECD, the second activity of DeSeCo was an analysis of existing theoretical and conceptual approaches to the notion of competence. This activity resulted in the report 'Concepts of Competence' by F. E. Weinert. (Weinert 1999) that is examined in Chapter 5.

The third DeSeCo activity, undertaken in 1999, was the production of five reports on competence by experts from different disciplines. The experts were asked to nominate competencies and to justify their approach theoretically, taking into account any available empirical research and providing evidence for the importance of the proposed set of competencies.

The DeSeCo Model of Competence

Figure 2 below was developed by Rychen and Salganik on the basis of the papers by Weinert and the other experts (Rychen and Salganik 2001). The underlying hypothesis for the model is that any designation of a particular set of key competencies is the result of multiple factors (Rychen and Salganik 2000 p.5).

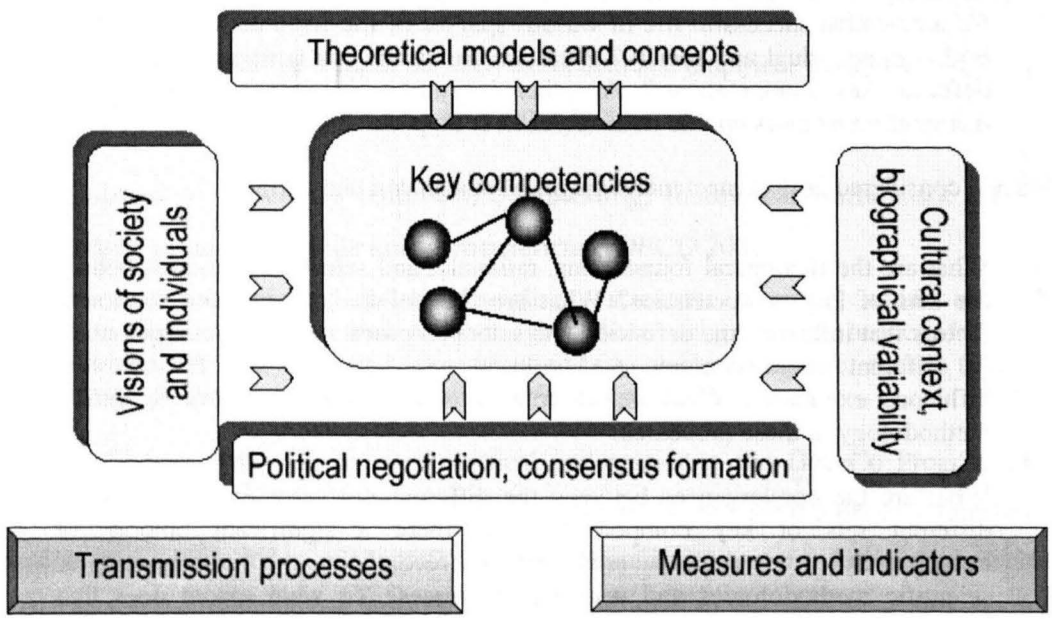


Figure 2 The DeSeCo Process of Defining and Selecting Key Competencies

According to Rychen and Salganik, at the most fundamental level, the underlying vision of the world, including assumptions about society and individuals, shapes the identification of key competencies. Different theoretical and conceptual perspectives also have a profound

effect on which competencies are identified as key as well as how the problem is approached. Factors such as culture, gender, age, and social status influence the forms that abstract key competencies take in specific contexts. Finally, the definition and selection of key competencies is a result not only of scientific analysis but also of a political negotiation process and consensus formation. According to Rychen and Salganik, each of these aspects should be taken into account when considering how key competencies are transmitted and developed, and when constructing and interpreting indicators.

Curiously, the DeSeCo model set out to take a rigorously scientific approach based on theory and empirical evidence, but it seems to end up giving remarkable recognition to political negotiation and consensus formation. A sketch below of the various papers on which this model drew gives suggestions about the promptings of these conclusions.

Weinert's Concept of Competence

The first substantive paper of DeSeCo entitled 'Concepts of Competence' by F. E. Weinert (April 1999) is by far the most salient, substantial and significant on those produced by DeSeCo. At the time of producing the paper, Weinert was researcher at the Max Planck Institute for Psychological Research. His paper reviewed the many 'scientific' uses of the terms 'competence' and 'competencies', and he identified the following broad range of meanings for the terms:

- general cognitive competencies;
- specialized cognitive competencies;
- the Competence-Performance model;
- modification of the Competence-Performance model;
- cognitive competence and motivation action tendencies;
- objective and subjective competence concepts;
- action competence;
- key competencies; and
- meta-competencies. (p.1)

The substance and conclusions of Weinert will be examined in Chapter 5 because this work is the most comprehensive and important attempt to deal with the notion of competence and competencies. At this point, however, it is worth noting that Weinert could find no consensually agreed core definition of competence. Similarly he did not see the identification of key competencies as resolvable in theoretical or empirical terms, and he seems to conclude that identification of key competencies is fundamentally a political negotiation and the formation of a consensus.

A bald summary (which will be justified in detail below) of the results of Weinert's analysis amounts to the conclusion that key competencies should be defined in

functional terms involving political negotiation and consensus. Thus the project set out to explore the theoretical and conceptual underpinnings of key competencies, but it implicitly concluded that the determination of competencies is not a matter of theory and research but rather a matter of political negotiation and consensus. It is in the spirit of Weinert's work that Rychen and Salganik note that 'concepts are generally socially constructed notions that facilitate the understanding of reality while also constructing it'. Ryochen and Salganik see the term key competencies as very fashionable in social policy discourse, but to them the term has very vague meanings. Ryochen and Salganik conclude that there is no broadly accepted definition or unifying theory of the notion of competence. Various 'scientific perspectives and ideological viewpoints' understand the term differently, and it is used differently according to different scientific and political objectives. Ryochen and Salganik conclude that DeSeCo should adopt a 'pragmatic conceptual approach, limiting the use of the concept with criteria which are more or less explicit, plausible, and scientifically acceptable' (p.8).

Weinert's Functional Approach to Competence

DeSeCo determined to take the 'functional approach' to the definition of key competencies recommended by Weinert. According to Weinert, the functional approach is concerned with what has to be done and it contrasts with the 'conditional approach' that is focused on internal mental structures that underpin cognitively defined abilities and skills. Under the heading that competencies are broader than knowledge and skills, Ryochen and Salganik describe the functional approach:

> DeSeCo focuses on a functional approach, which places complex demands facing individuals at the forefront of the concept of competence. According to this viewpoint, competencies are structured around demands and tasks. Fulfilling complex demands and tasks requires not only knowledge and skills but also involves strategies and routines needed to apply the knowledge and skills, as well as appropriate emotions and attitudes, and effective management of these components. Thus, the notion of competencies encompasses cognitive but also motivational, ethical, social, and behavioural components. It combines stable traits, learning outcomes (e.g., knowledge and skills), belief-value systems, habits, and other psychological features. In this view, basic reading, writing and calculating are skills that are critical components of numerous competencies (Rychen p.8).

The nature of this approach and the issues it raises will be considered in Chapter 5, but it is worth noting here that this definition of competence and competencies gives priority or emphasis to the 'complex demands facing individuals', thus placing what has to be done at the forefront. In this sense DeSeCo is like the other conceptualisations above. They generally define work-related skills in terms of what has to be done, and they offer little or no

recognition of the cognitive capacities that underpin the performance of different tasks. Thus, while DeSeCo takes a completely different approach and heads in a quite different direction to the other work-related skills projects, it ends up in much the same position as they do when it concludes that competencies should be defined in terms of kinds of tasks. Where the other work-related skills projects assumed that the definition of competencies was a kind of negotiation between 'stakeholders', DeSeCo came to the same conclusion after extensive conceptual and theoretical review.

The Opinions of the DeSeCo Experts on Competence

The opinions of experts in various scientific disciplines (e.g., sociology, economics, philosophy, anthropology, psychology, etc.) constitute the third, and primary, activity of the DeSeCo Program. A small number of 'scientific experts' were commissioned to prepare reports presenting their proposals for a set of key competencies grounded in their theoretical perspectives (Rychen and Salganik 2003). The results of this work are summarised in Table 4.

The experts commissioned by the OECD structured their approaches to identifying key competencies around different organizing ideas, central questions, and conceptual frameworks. One of the invited experts, the anthropologist Jack Goody, rejected engaging in a decontextualised discussion of key competencies on the grounds that theory must always be considered in the context of practice. Recognising that there may be some very general qualities required by modern life, Goody focused on the intractability of specifying key competencies at a level that can span cultures, social contexts, and individuals and at a level that would also be useful for developing measures. He also cautioned against limiting the work to developed countries because it is bound to be used in a larger context and have a negative, homogenising effect.

With the exception of Goody, the DeSeCo experts each proposed a set of key competencies needed by the individual for what the authors define as success in terms of their overall approaches. The philosophers, Canto-Sperber and Dupuy, identified the competencies needed by the individual for a good and meaningful life. Haste based her competencies on psychological models of how individuals interact with their environment. Levy and Murnane, as economists, detailed those competencies needed for success in today's labour market, while the sociologist Perrenoud identified key competencies as those needed for the autonomous actor to successfully defend his or her rights within society.

Table 4 The Key Competencies Identified by Four Expert Reports of DeSeCo

The philosophers **Monique Canto-Sperber & Jean-Pierre Dupuy**	**The psychologist** **Helen Haste**	**The economists** **Frank Levy and Richard Murnane**	**The sociologist** **Philippe Perrenoud**
Five clusters of competencies needed for a good life: Coping with complexity (recognizing patterns); Perceptive competencies (discriminating between relevant and irrelevant features); Normative competencies (choosing the appropriate means to reach a given end, appreciating various possibilities offered, making moral judgments and applying them); Cooperative competencies (cooperating with others, trusting others, taking the role of the other); and Narrative competencies (making sense of what happens in life to oneself and others, describing the world and one's own real and desirable place in it).	Key competencies: to: adaptively assimilate changing technologies into social practice (technological competence); deal with ambiguity and diversity; find and sustain community links; manage motivation and emotion; and desire; focus on morality, responsibility, and citizenship. (agency and responsibility)	Key Competencies to: reading and mathematical skills (not only for their instrumental use but as the basis for lifelong learning); oral and written communication abilities; skills to work productively in different social groups; emotional intelligence and related abilities to co-operate well with other people; and familiarity with information technology.	Key Competencies to: identify, evaluate, and defend one's resources, rights and limits; form and conduct projects and develop strategies, individually and collectively; analyze situations and relationships; to co-operate, act in synergy and share leadership; build and operate democratic organizations and systems of collective action; manage and resolve conflicts; understand, apply, and elaborate rules; and construct negotiated orders beyond cultural differences.

Adapted from Rychen and Salagnik 2000

The Outcomes of DeSeCo

In reporting to the General Assembly of International Indicators of Educational Systems project of the OECD in 2000, Rychen and Salganik identified the following normative positions as a vision of society and individuals arising from DeSeCo:

- Key competencies are consistent with the principles of human rights and democratic values;
- Key competencies give individuals the capacity for a good, successful life; and
- Key competencies are not incompatible with social and individual diversity. (p. 6-7)

Rychen and Salganik identified considerations and arguments that transcend the 'heterogeneity of these discipline-oriented approaches to competencies' (p.11). They identified four key elements, the three processes and the five dimensions that compose into the three generic key competencies shown in Table 3. These three key competencies aimed to encompass many of the key competencies identified in the five expert reports.

Table 5 Aspects of DeSeCo Key Competencies

Conceptual elements of key competencies	Multidimensional processes compositing key competencies	Five dimensions of key competencies	Three generic key competencies
Key Competencies: are multifunctional are transversal across social fields refer to a higher order of mental complexity are multidimensional	Multiple dimensions representing mental processes composed of 'know-how' analytical, critical and communication skills common sense	**Coping with complexity** Recognizing and analyzing patterns, establishing analogies between experienced situations and new ones. **Perceptive dimension** Perceiving situations, discriminating between relevant and irrelevant features. **Normative dimension** Choosing appropriate means in order to reach given ends, appreciating various possibilities offered, making judgments and applying them **Cooperative dimension** Developing social-orientation, trusting other people, listening and understanding others' positions. **Narrative dimension** Making sense of what happens in life to oneself and others, seeing and describing the world and one's real and desirable place in it	Commonality in discipline-oriented approaches Acting autonomously and reflectively Using tools interactively Joining and functioning in socially heterogeneous groups.

(Adapted from Rychen and Salagnik 2000)

In their overview of the DeSeCo project when three quarters of the way through, Rychen and Salganik described a developmental or iterative process of theoretical and empirical work (including the development of measurement methodologies) as the key to future success in this field:

> Yet, to obtain such information, considerable effort is needed in the areas of scientific inquiry and political discussion. We need a better understanding of how different social fields operate and how this affects demands on individuals and society. Further, we need a better understanding of the relationship between general social conditions (such as globalization and new technology) and the demands on individuals and society. An iterative process between theoretical and empirical work (including the development of measurement methodologies) is key to future success in this field. Recognizing that defining and selecting key competencies is ultimately the result of a political negotiation process, in which researchers can only be partners among others, we encourage within the DeSeCo Program the dialogue between scholars, representatives of the economic world and the social sphere, as well as policy makers. As we continue to develop our ideas about key competencies, we need to work towards a conceptualization that is both theoretically grounded and relevant to the policy world (Rychen and Salagnik p.14).

So DeSeCo concludes that the definition and selection of competencies is ultimately a political negotiation: while the DeSeCo competencies aspire to be consistent and coherent theoretically they must also be politically meaningful and relevant.

A Commentary on DeSeCo

The DeSeCo project was the most adventurous and sophisticated of the generic and work-related skills projects. For all its concern with theory and conceptualisation, DeSeCo comes to the conclusion that the definition and selection of competencies is in a fundamental way pragmatic and political. DeSeCo approached the notion of theoretical grounding of competencies in a curious fashion. It assumed that by asking members of a range of different disciplines to address the same issue a consensus or synthesis could be attained. It will be argued in Chapter 5 that this approach, as taken by Weinert and in the notion of agreement between different disciplines, is very optimistic.

Psychometrics and cognitive psychology are the disciplines concerned with theorising performance and abilities, but they are not represented among the DeSeCo experts, and although Weinert is a psychological researcher concerned with performance, he is clearly an opponent of the work in psychometrics and cognitive psychology that is concerned with general abilities and generic skills.

The perspectives of the philosophers, the economists, the sociologist, the anthropologist and the social psychologist offer little coherent theory about performance and abilities. There

seems to be little that can be usefully generalised from the perspectives of the discipline experts, and the attempt of Ryochen and Salganik to synthesise this work is not very closely related to the perspectives of the experts.

The stipulations under the heading 'Conceptual elements of key competencies' in Table 3 offer the potentially tautological claim that key competencies are multifunctional and transversal. It is asserted that they are a 'higher order of mental complexity' but it is hard to see what this meaningfully adds or excludes. One of the elements of competence is multidimensionality which is the notion that mental processes include 'know how', analytical, critical and communication skills, and commonsense. These multiple dimensions of cognitive processing seem to be envisaged as crossed by dimensions for complexity, perception, normativeness, cooperativeness and narrative. It seems to be implied that a kind of two dimensional matrix of mental process and other dimensions in some sense to add up to the three key competencies of:

- acting autonomously and reflectively;
- using tools interactively; and
- joining and functioning in socially heterogeneous groups.

These notions do not evidently arise from the texts on which they are based, nor do those texts offer a convincing basis for the three key competencies. It turns out that the DeSeCo key competencies are the least practical or useful of any of the conceptions reviewed here, and they offer the least practical basis for developing indicators for the international comparison of educational outcomes.

The VCAB Student Profile of Work-related Capabilities

Even though the scope of each of the conceptions outlined to date is national and international, it is worth mentioning a more parochial conception, the Student Profile of the Victorian Board of Studies, that had brief life in the early 1990s. This conception is interesting because it is cross-curricular and work-related and because it deals with personal characteristics. The VBOS Student Profile is also interesting because it was designed by a committee of practicing teachers as an assessment framework and it was, briefly, implemented as such.

The VCE Student Profile was developed as a part of the Victorian Certificate of Education and implemented in 1992 (VCAB 1990; VCAB 1994). It was to accompany the statement of results leading to the VCE and was intended to complement the Common Assessment Task grades given in a student's Statement of Results:

> Additional information about students' work-related capabilities and other achievements in Units 3 and 4 will be reported by schools in the form of a cross-curriculum statement presented in a common format designed to convey information about student capabilities which is useful for employers and for tertiary selection (p.1).

Three principal elements were assessed by the Profile:

- cross-curriculum capabilities;
- demonstrated significant improvement; and
- 'other comments'.

The work-related capabilities assessed by the Profile were:

- initiative;
- self-management;
- cooperative work;
- adaptability;
- reflection/evaluation; and
- communication.

Assessment of these capabilities was based on each study-teacher's observation of a student's work during year 12. Study-teachers were assisted in making judgements by the provision of a definition of each capability, followed by model descriptions of High, Medium and Low performance on that capability. Each study-teacher provided assessments which were summarised and reported in the form of a cross-curriculum description for each capability .

'Communication' is the only curriculum-related skill in the work-related capabilities of the Profile. Rather than being cognitive and academic, most of the Profile capabilities are social and personal skills and hence involve matters of attitude or personality. The Profile was able to draw on the involvement of students in extra-curricular school activities and non-school activities, and it offered an opportunity to assess and report on a student's social and personal skills and attitudes.

The Student Profile is the only system-wide scheme of assessment based on the global, impression judgements of teachers assessing cross-curricular capabilities and attitudes to produce a whole-school assessment in Australia. Given that the capabilities are similar in nature to some of the Key Competencies (they were developed at about the same time) in that they are a mixture of cognitive and attitudinal attributes, the Profile is of particular interest as a comparison with the Mayer Key Competencies, and the proposals to assess them.

At this point it would be appropriate to review the different conceptions of generic work-related skills presented in Table 1.

A Comparison of the Key Competencies and Other Generic, Work-related Constructs

The Secretary's Commission on Achieving Necessary Skills (SCANS) construct is significantly more elaborate than the Mayer Key Competencies. The Key Competencies were challenged during their development and subsequently on the issue of the 'foundation knowledge' that is supposed to 'underpin competence'. The SCANS conception does not present a conception of foundation knowledge. Rather, it views the foundation of other competencies as a matter of basic literacy, oracy and numeracy, and also defines this foundation in terms of kinds of cognition (Thinking Skills) and affective or attitudinal constructs (Personal Skills). The comparison of these foundation skills with the Key Competencies shows how broad the single Key Competency of Communication is in the Mayer scheme. Similarly it shows how many kinds of thinking have been included in the Problem solving Key Competency. Unlike the Key Competencies which explicitly eschewed values and attitudes, the SCANS construct includes a range of personal skills and characteristics. The SCANS conception envisages applying these foundation elements to five competencies: the specific context of technology, the more general contexts of resources, information and systems, and the interpersonal attributes involved in teamwork. The foundation elements of the SCANS stand as the enabling, generic skills (literacy, oracy and numeracy), general thinking skills and personal attributes. The interpersonal competencies are a natural extension of the foundational personal qualities, and the technology competency is a fairly tangible (if not unproblematic) construct. The 'resources and systems' competencies seem very vague, and it was wise of the Mayer Committee to eschew such vague notions as these.

The SCANS conceptualisation is more definite about what it takes to be the foundation of its 'know-how' than the Mayer conceptualisation, and it is clear about its focus on personal skills and qualities. However by comparing SCANS 'know how' with the Key Competencies we can see that the emphasis on activity in the Mayer conception makes them more specific and concrete.

The British Key Skills have a basis of literacy, oracy and numeracy. The personal skills of the earlier core skills conception have been turned into working with others in the same way as happened to the personal skills of the Finn conception in their transformation into the Key Competencies. Improving one's own learning is a rather vague notion that would be difficult to differentiate from the ability to learn, or the speed with which an individual learns. In the transition from Core Skills to Key Skills, the British conceptions have moved toward the Key Competencies, but the Australian Key Competencies are more concrete than the British Key Skills.

The South African cross field outcomes cover literacy, oracy and numeracy, make references to knowledge of science and technology. The outcomes also include the basic cognitive activities of dealing with information and problem solving, the personal qualities involved in team work, and the notion of the world as an interrelated system. Unlike the other conceptions, the South African outcomes make specific reference to science and to the world as a set of systems.

The Canadian Employability Skills Profile includes the basic academic skills, self-management and teamwork. The Canadian Essential skills are also literacy, oracy and numeracy, generalised thinking skills, working with others, computer use and the general notion of continuous learning. The most recent conception of employability skills in Canada involves the fundamental skills of literacy, oracy and numeracy with skills in managing information and thinking. Positive attitudes are picked up in the personal management skills, and teamwork skills also entail planning and organising.

The OECD key competences are the most global and abstract of the conceptions, and as such they seem to be the least specific or practical. There is no relationship to the educational curriculum or to cross-curricular skills in the DeSeCo competencies. It is not possible to be more generalised than the notion of 'acting autonomously and reflectively'. The notion of autonomy in the DeSeCo conception is an ideal or aspiration. It is impossible to imagine as a more or less specific activity, or even a generalised function. It is not a function. Reflection can mean thinking in general, but it has overtones of stance, attitude and personality. The notion of reflection takes thinking a long way from a more or less specific action or even a generalised function. The notion of joining and functioning in socially heterogeneous groups is a description of the human condition rather than an ability or competency. Using tools interactively is a kind of function, but it is at the level of defining the characteristics of our species rather than at the level of a useful competency. So as a group, the DeSeCo competencies are the least specific and the least functional of all the conceptions considered here. They are the least realistic and offer least guidance in developing an assessment of any of the work-related notions considered here.

Leaving the DeSeCo competencies aside, there is a significant degree of similarity and some diversity in these conceptions. They can be readily compared in terms of:

- basic, enabling or foundation skills of literacy, oracy and numeracy;
- cognitive, research and problem solving skills;
- areas, domains or kinds of knowledge and skill, such as technology and systems and technology; and

- personal attributes, from working with others through to adaptability.

On one hand all of these conceptions pick up the basic enabling skills, and on the other they all include reference to the cognitive, research and problem solving skills. All refer to technology or the more specific information technology, but the scholastic content areas are barely mentioned in any of them. Science is only referred to in the South African outcomes, and learning about culture and society, that might have been included as Cultural Understanding in the Mayer Key Competencies, is not to be found in the other conceptions. The most distinctive feature of these conceptions is the prominence given to personal attributes, the kinds of matters that involve attitudes and personality. In comparison with the other conceptions, the Mayer Key Competencies seem the most cautious in focussing on the assessment of cognition rather than the assessment of attitudes. Perhaps this caution in the Mayer conception was shaped by the radical nature of the Mayer assessment proposals. In conclusion, it seems that the Mayer proposals offer the most realistic (and cautious) basis of any of these schemes on which to develop an assessment.

The Assessment and Reporting Recommendations of the Mayer Committee

One of the areas in which the Mayer proposal contrasted most markedly with the proposals in other countries was in the breadth, specificity and audacity of its assessment proposals. The Committee aimed for 'national consistency in assessment and reporting' for the first time in Australia on the basis of Key Competencies. They claimed this national consistency should be facilitated to obtain:

- portability of the assessments;
- consistency of credit transfer; and
- ease of interpretation by users.

The Mayer Report notes that it is sometimes assumed that national consistency can only be achieved by the use of uniform methods, and even common instruments, for assessment; but the Committee stated that such an approach would be neither feasible nor appropriate for assessing Key Competencies.

The Committee argued that because there are diverse programs, settings and modes of program delivery, the Key Competencies must be assessed with a variety of methods. They also argued that a variety of assessment methods is demanded by differences between the competencies themselves. According to the Mayer Report, no single assessment method can be regarded as most appropriate for assessing all the different Key Competencies.

Consequently the Committee recommended pursuing national consistency through nationally-agreed principles, rather than common assessment instruments or methods, because of the need to accommodate the diverse range of programs and settings through which post-compulsory education and entry-level vocational training are delivered. The Committee did not formally outline an assessment system, but it offered the principles for assessment that aimed to give both consistency of approach and the flexibility necessary to accommodate diverse programs and settings.

Most of these principles for assessment offered by the Mayer Committee are unexceptionable, but they are not particularly concrete or informative (see Appendix 1). They do not offer much guidance for design or implementation of an actual assessment, and some of the principles, and the assumptions on which they are based, have been subject to the criticisms mentioned earlier in this discussion. The procedures suggested by these principles would provide a Record of Performance, including a performance level on each Competency to any student exiting a Year 11, Year 12 or equivalent course. The assessment would be made and issued to the student by the education provider, and the aggregate information on all students would be collected into databases that could be sampled for system level accountability monitoring. Overall, this proposal is an exceptionally ambitious scheme for a national assessment, and not surprisingly it has been subject to a number of criticisms. Concerns have been expressed about the context-dependent nature of performance, and uncertainty about the number of contexts needed for reliable assessment has been a cause of critical comment. More broadly, the reliability of provider-based judgements as an assessment method has been challenged.

The Assessment Assumptions of the Mayer Committee

One can disinter the following from the assessment principles of the Mayer Committee.

The assessment of Key Competencies:

- is to be contextualised in current courses and procedures, and it need not entail new assessment activities;
- need not be based on specific tasks and could be integrated into other assessments, but the assessment would be a matter of separate reporting;
- would be based on absolute standards (three described levels of performance) and the assessment would not be norm referenced;
- would not be subject-based, rather it would be across a whole course or program and it would entail some process of aggregation within a course or institution; and
- would be comparable between, and could be aggregated across, providers.

These assumptions pose formidable pragmatic difficulties to even the most enthusiastic of systems intending to implement the scheme.

We will now turn to look at various responses to these proposals of the Mayer Committee.

3

The Generic Skills Debate in Australia

Since the publication of the Mayer Report a number of reservations and doubts have been expressed about the assumptions made and the proposals offered by the Committee. The views of some of the proponents the Key Competencies will be briefly outlined below before turning to the reservations and doubts.

While there are good reasons for questioning the soundness of the Mayer proposals, it should also be noted that the Key Competencies represented a politely phrased but nonetheless real challenge to the vested interests and power structures of education. As a result, the Key Competencies excited a good deal of direct and indirect criticism of a political kind from the educational community.

Some of the more fundamentally political criticism of the Key Competencies proposals expressed itself as a theory-based argument about learning and assessment, and it is this argument that will be canvassed here. The Key Competencies debate is unusual because it was often phrased in terms of a view of ability and what is assessable, but there is a sense in which these theory-based arguments were not unrelated to the vested interests of the critics.

Arguments for the Key Competencies

Anne Borthwick, the secretary of the Mayer Committee, commended the 'broad definition of competence - characterised by applied focus and cross-curricular orientation' of the Mayer proposals She described the proposals as a progressive 'move away from the traditional narrow approach of norm-referenced assessment or standardisation' and argued that the Key Competencies were concerned with standards based on criterion referenced assessment procedures. According to Borthwick, they would involve the 'use of subject profiles for reporting achievement'. She describes the Key Competencies as concerned with 'standards not standardisation' (Borthwick 1993 p.244).

Laurie Carmichael did not agree with those who feared the Key Competencies were narrowly vocational (Carmichael 1993). According to Carmichael changes in the economy, work practices and technology were leading to a convergence of general and vocational education:

> It becomes increasingly difficult to distinguish between what is general and what is vocational education for modern employment purposes in the modern mature labour market (Carmichael 1993).

In reviewing the British National Vocational Qualification he described it as promoting 'broad based competence', and noting that in the future 'work will become more and more professional and the gap between humanities and instrumentalist pursuits will become less and less' (p.20).

Simon Marginson (Marginson 1992) described the idea of competencies as an advance, in that it combined the 'older ideas of skills and knowledge' (p.6). Like Carmichael, Marginson saw competencies as bringing about a convergence of general and vocational education so as to make the general more vocational and specific and the specifically vocational more general:

> The idea of competencies combines the older ideas of skills and knowledge. This is an advance. The only social distinction between the general and the vocational was a skills/knowledge split; originally, a division between those who used their hands (skill) and those who used their heads (knowledge). "Competencies" means that people who have skills also have knowledge and vice versa (p.6).

Colin Ducker (Ducker 1993) from a VET perspective described the Key Competencies as potentially more challenging to the status quo in vocational education than to general education:

> The key competencies present some significant challenges to the training curriculum not only in the area of training curriculum design but also in related areas such as assessment (p.74).

According to Ducker, the Key Competencies align more closely with broad notions of competence than with the 'narrow task based ideas of the past. Indeed, the key competencies are well placed to foster the growth towards broader notions of competence' (p.74).

The Key Competencies Pilot Phase Review (MCEETYA 1996) described the Key Competencies as consistent with and enhancing general education:

> The Key Competencies are a set of eight very general competencies, the main purpose of which is to develop the relevance of general education and promote flexible and adaptable learning in later life, especially, but not exclusively, in the workplace. They do not displace other goals of education, but extend and complement them. The Key Competencies are not intended to be the whole curriculum but a means of extending the relevance of the curriculum. They are consistent with developments in education of a similar kind over many decades, but they crystallise and give greater explicitness and focus to the idea of relevance than many earlier ideas did (p.1).

Criticisms of the Key Competencies

On the other hand, the Key Competencies have been criticised on the grounds that:

- generic or transdomainal skills and hence the Key Competencies do not exist;
- competence is knowledge-based and context-specific;
- generic skills are unassessable; and
- the Key Competencies are narrowly vocational and they undermine vocational education.

It seemed to some that the Mayer Committee made untenable assumptions about the nature of skill and the nature of learning (Collins 1993). To critics the Key Competencies were based on a mistaken view of competence in that they overlooked the importance of specific knowledge and skills in competent performance. According to this competence-as-expertise view, competence is knowledge-dependent and expertise is knowledge-rich. Evidence for the transference of training from one context to another is poor and hence the notion of generic or transdomainal skills is vacuous. In a cognate argument it was claimed that the Key Competencies minimised or ignored the fundamental importance of discipline-based knowledge and technical knowledge and skills in education and training. It was feared that the Key Competencies would replace academic competence (disciplines, objectivity, truth) with operational competence as know-how and skills. This operationalism was criticised because it would replace what students understand with what students can do (Barnett 1994).

Consequential objections were raised to the Key Competencies on the basis that they are:

- merely instrumental and narrowly vocational;
- undermining general education;
- reducing general education to job training; and
- diluting vocational education.

Pragmatic objections were raised to the Key Competencies because:

- they are new, foreign and unassessable constructs;
- performance cannot be generalised because it is inconsistent and domain-specific;
- competence and expertise are knowledge-dependent; and
- teachers cannot make judgements about such constructs.

The Key Competencies were criticised as yet another burden to an over-crowded curriculum that would dilute the rigor of education. They would be costly and impractical to implement, and they would be time consuming and an added workload for teachers and students.

Key Competencies were seen as at odds with the structure and organisation of schools. The cross-curricular nature of the Key Competencies could not be accommodated in conventional curriculum, and they would distract from and undermine the disciplined-based learning of knowledge. It was contended by some that educational institutions should be concerned with knowledge and skills and that adaptation to the workplace should be a matter of on-the-job training. Overall it seemed to such critics that the Key Competencies proposals were unrealistically ambitious and impractical.

Such pragmatic issues are weighty considerations, and it will be seen when looking at the use of the Key Competencies in Chapters 8 and 9 that such issues had to be considered in any realistic attempt to implement the Key Competencies. It sometimes seemed that the advocates of the Key Competencies who hailed them as an educational revolution did little to encourage the adoption of them by education systems. On the other hand, the work reported in chapter 9 aimed to explore the assessment of the Key Competencies with minimal investment of time and energy by teachers and with minimal change to what schools usually do.

The Mayer Committee was criticised for lacking a clear theoretical position and for not drawing on research in formulating policies. These critics taxed the Mayer Committee with being atheoretical and ignorantly or irresponsibly formulating proposals that are at odds with 'what research shows' and 'what we know about the mind or brain' (Collins 1996).

This chapter will outline these criticisms of the Mayer proposals and Chapter 6 will examine 'what research shows' and make some comments on 'what we know about the mind or brain'.

Theoretical and Research-based Objections to the Key Competencies

Some of the earliest disquiet about 'competencies' was expressed by the then Director of the ACER, Barry McGaw. He had been involved in the Quality Education Review Committee described in the preceding chapter, and he responded to the early interest in generic skills by noting in the Annual Report of ACER the benefits of an interest in competence (McGaw 1992):

> The great benefit of this approach is that it carries a strong emphasis on the outcomes of education. Course constructors and teachers have to be clear about what competencies they intend their students to acquire and then to concern themselves with whether their students acquire them. That is not a bad thing since teachers can be tempted to see their responsibilities ending with teaching and not extending to students' learning. Along with the benefit of attention to skill development, however, are some significant risks (p.10).

He concluded his discussion of the issues with the following reservations:

> (1) Attention to generic competencies at the school level may divert attention from specific subject competence. Research on the nature and acquisition of expertise makes clear that the manner in which experts represent and solve problems is largely specific to their domain of expertise and that it is highly dependent on knowledge of the domain.
>
> (2) The link between some general intellectual competencies and specific courses of instruction is likely to be weak. How, for example, might a historian define the link between the specific competencies of a professional historian and the content of a particular history course?
>
> (3) Some of the most important competencies may be the most difficult to define in ways that make the task of assessing them feasible. Related to this is the risk that it may then lead to greater emphasis on those competencies that can be most easily defined. This is just a new version of the old complaint that courses are too often reduced to only those things that an external examiner can assess in a three hour written examination. New courses would be reduced to a concern with only those competencies that can be most easily specified and assessed (p.10).

McGaw concluded his discussion with a warning about the dangers of competencies trivialising education:

> In summary then, Australia is in the midst of some radical rethinking about education. Some of it is driven by a continuing growth in demand for education; some of it by concern that education is insufficiently linked to workforce requirements. The benefits are likely to be new thinking about the options available in the post compulsory years, serious attempts to define multiple pathways and to facilitate students' movement through and between them, and a healthy focus on the benefits that students should obtain at each stage of their education.
>
> The risks are that the hierarchical arrangements of post-school institutions that have been strengthened through the almost unfettered growth of universities will be too strong for a genuinely respectable technical education stream to emerge to prepare highly valued people with highly valued technical skills; and that the new attempts to focus on the development of specific competencies will trivialise the educational process because competencies will be on the one hand, too general to be useful or, on the other, specified in ways that encourage people to lose sight of the integration of competencies that characterises the performance of experts (p.11).

These substantial and significant warnings were offered before the Mayer Committee produced its report. As it transpired, the Mayer proposals did not envisage an alternate curriculum that would distract from or marginalise the academic or technical content of the education.

The argument about the nature of expertise (and competence) will be examined below, but it might be noted at this point that assessing generic skills is not about assessing a degree of expertise (or competence in the sense of expertise). Generic skills assessment is about prediction of the ability to learn. In other words, it is about the aptitude to move with facility

from being a novice to being an expert. Becoming an expert may well be fundamentally based on the acquisition of specific knowledge and skills, but not every one can reach the level of an expert. Individuals move from novice to expert at very different speeds, and some expertise is facile and some is laboured. A generic skills assessment is concerned with predicting the ability to learn with facility and to reach a high level of expertise with a cost-effective investment of resources. A generic skills assessment does not claim to account for the change from novice to expert.

Reservations of the kind offered by McGaw surfaced in various reviews of education in the 1990s. Piper noted that:

> In general, the key competencies have had a lukewarm reception in the educational community. The major criticisms have been the excessive focus on employability, the instrumental view of education as an arm of economic policy, and a questioning of the validity of the concept of generic competencies, divorced from specific content (Piper 1997).

Piper goes on to instance the reservations expressed in the 1994 review of school curriculum in Queensland.

> Thus the report of the Review of the Queensland School Curriculum (Wiltshire, 1994: 101):
>
> > The key competencies which have evolved at the national level began their life as employment-related competencies or skills and not educational competencies. This remains fundamentally the issue and, whilst educators might well agree with the notion that certain competencies are vital for the workplace and their identification can assist in the transition to employment, their relationship to schooling in general and curriculum in particular is not simply a matter of instantaneous definition, location, isolation, and measurement in a schooling system (p.51).

The views presented by McGaw at the beginning of the decade shaped the recommendation of the major review he chaired into the Higher School Certificate in New South Wales (McGaw 1996). That review resulted in one of the most explicit and the most important rejections of the Mayer proposals:

> **Reporting key competencies**
>
> The Government supports Professor McGaw's advice that there be no central reporting of students' performance on key competencies. Professor McGaw's comments, along with the findings of the New South Wales Key Competencies Pilot Project, indicate that it is best to focus on the key competencies within the context of the curriculum and not develop a separate system-wide approach to assessing and reporting on key competency achievement. Key competencies can provide a useful language for describing attributes that are valued by teachers, trainers, students, parents and employers.

> These attributes have been variously described and used in school reports, references, job advertisements and recruitment practices. Reporting that provides additional information on student achievement in these areas is broadly supported.
>
> The Government agrees with Professor McGaw that there should be no central reporting on key competencies, but that schools should have the option of providing reports. Reporting at school level provides better opportunities for including contextually rich evidence of students' achievements of key competencies.
>
> The Board of Studies, Department of School Education and TAFE New South Wales will provide support to schools and colleges issuing their students with reports on achievements of those aspects of the Higher School Certificate that are consistent with the key competencies (p.26).

This policy pronouncement in 1996 seemed to presage the end of the Mayer vision of a national assessment of Key Competencies. A change in the federal government in 1996 seemed to confirm this outcome. There had been a good deal of debate about the issues between the expression of McGaw's reservations in the early 1990s and the results of the NSW review in 1996.

For instance, in an article entitled 'How could the `Generic Skillers' and `Competencies Buffs' be both right and wrong? a defence of organised knowledge', Craigie described the Key Competencies as at odds with 'cognitive science'(Craigie 1995):

> Whether human understanding and intellectualising are best served by general domain-independent skills and abilities, by domain-specific skills and abilities, or by a mixture of the two has been a recurring debate. Recent moves towards 'key competencies'; generic thinking skills programs, and broad curriculum statements only tenuously tied to underlying academic disciplines, whilst striking a useful note towards operationalising knowledge, seem to misunderstand in many ways the essential nature of knowledge, and also to misread or be unaware of recent developments in cognitive science, which stress the essential place of the knowledge base in developing cognitive competence (p.169).

Talking about human capacities to comprehend the world, Craigie argues that although generic skills are important, it is essential to comprehend 'the logically arranged schemas of concepts and generalisations in what have traditionally been called academic disciplines' (p.170). He sees generic skills as playing a complementary but not a dominant role in general education.

The Arguments of Collins and Stanley in the Competencies Debate

The most explicit and best formed criticisms of the different aspects of the competencies movement was published in 1993 in a book entitled Competencies: The Competencies Debate in Australian Education and Training (Collins 1993) edited by the academic and

researcher Cherry Collins. Collins found the notion of an economic impetus behind the competencies movement objectionable as such:

> ... much of the competencies agenda is openly embedded in an economic vision which sees a need to harness our educational institutions to economic ends (p.12).

She saw both specific and generic competencies as growing from 'behaviourist roots':

> The Mayer Committee, for all its care to move beyond behaviourism in its definition of competence, stayed mired in behaviourist fundamentalism in its envisaging of the monitoring of levels of competence: its report assumes a stepped progression of pre-specified outcomes, seen as cross-curricular and as checkable (yes/no) from simple performance (p.8).

She described the Key Competencies as a 'skewed list' because they made no reference to 'critical or caring thinking'. She summarises with particular approval the critique offered of the Key Competencies by Gordon Stanley, a contributor to the book, in the following terms:

> The final section contains three critiques. The first of these, by Gordon Stanley of the Western Australian Office of Higher Education, looks at psychological research, particularly that on human cognition, and applies it to the competencies debate. He is particularly skeptical about the idea of cross-curricular competencies, the assumption on which the key competencies agenda is based. There is extensive research evidence that human capabilities are domain specific. For example, problem solving capability in mathematics is entirely unrelated to problem solving capability in the inter-personal domain. 'Problem solving' thus cannot be defined, by committee fiat, as a single, cross-curricular competency. Further, he argues, the more expert one is, the more one's capabilities are integrated into the domain's knowledge base and the less one's competencies are transferable. Gordon is also skeptical about the possibility of reliably assessing competence from performances, as proposed in all initiatives. He concludes that there is 'no substitute for the building up of knowledge bases in specific domains' (p.10).

Collins made a more overt and detailed critique of the Key Competencies (and the National Curriculum Profiles) in a publication entitled 'What Teachers Need to Know? The Competencies Debate' (Collins 1996). According to Collins, the Mayer Committee had ignored (through ignorance or disingenuousness) fundamental questions about the nature and purpose of the curriculum. She questioned the fundamental assumptions about knowledge made by the Mayer Committee that 'there are seven generic (listed by Mayer) human capabilities, which human beings can learn and then summon forth', and that 'these generic capabilities can be learned in one place for use in others, indeed learned in a few school subjects for use in practical situations in the employment world' and that 'persons can learn a generic capability in one setting and transfer it to another' (p.7).

Collins says that while the Mayer Committee made many unwarranted assumptions, she chooses to:

> ...focus simply on some empirically testable assumptions about human beings and generic capabilities. Whether these assumptions make sense is not a matter of debate but of empirical evidence and, therefore, particularly telling. Do Key Competencies, at this empirical level, make sense (p.7)?

Collins implies that these issues can and have been empirically determined in the negative:

> The generic capabilities of human beings are matters on which we have ninety years of research evidence. It was in the first decade of this century, in response to another group of wishful thinkers, that E. L. Thorndike first showed that *homo sapiens* don't learn in the ways that employers' dreams are made of. His wishful thinkers were not employers, as it happens, but Classicists who advocated the teaching of the Classics on the grounds that one learned therein generic mental capabilities which transferred to other aspects of life. What Thorndike showed, however, still applies (p.7).

Collins goes on to state that:

> This is not just the conclusion one must reach from Thorndike's early work on transfer of training, but the conclusion of almost continuous work on human cognition since, and particularly of recent work on the development of expertise (Stanley, 1993; Ericsson and Smith, 1991). Evidence from cognition research, and there has been a great deal, is that competence is field specific, not generic. It doesn't just transfer from one aspect of life to another (p.7).

Collins chooses the notion of problem solving, the usual target of critics of generic skills, to develop her case:

> Let us take problem-solving, one of the Key Competencies as an example because the issue becomes very clear when we do. What has solving the conceptual, aesthetic and skill problems of painting an art work got to do with solving the problems in an algebra set? People who are excellent at one of these are rarely good at the other. To talk of 'problem-solving' as a generic capability, a Key Competency that one can simply transfer from art to mathematics is simply a nonsense. People's problem-solving skills are field specific, not generic. They are embedded in knowledges, sensitivities, familiarities and understandings of particular fields and particular situations. The knowledges and skills required for solving the problems of, for example, classical piano performance, are a schemata specific to the field, not the sum of competencies generic across fields that are summoned together and applied to that field. Most people competent in classical piano performance are not even competent at jazz piano performance. They may be no good at all at wood carving, although both piano playing and wood carving involve the "Key Competence" of fine motor skills, not at bird call identification, though both require the "Key Competence' of skilled listening. If one wants young people to become competent, then one has to get well past the naiveté of the generic, cross-field way of thinking about it that Key Competencies advocates fantasised (p.7-8).

Collins concludes these comments by stating that

> Key Competencies are an example of the kind of common sense hocus pocus which gets thrown at education systems and which they get blamed for failing at, even though success is literally humanly impossible. We cannot teach problem-solving as a generic skill. At least we can't teach it to human beings. (p.8)

For Collins the Key Competencies are not plausible or feasible and they were a waste of time and money:

> While it is unarguable that we know that 'people's problem-solving skills are field specific, not generic', but all is not lost, we know some things about transfer.
>
> We do know some things about how to help students transfer some of their skills and knowledges to strange situations like new work environments. Some useful heuristics have been researched. But the key is teaching quite specific meta-cognitive skills and attitudes, not fantasising the learning of levels of generic capability (p.8).

Collins concludes that the Key Competencies (and the National Curriculum Profiles) are 'not credible in the light of current research evidence and debate (p.11).

Collins (like Craigie) dismisses the Key Competencies on the basis of 'what research shows' and 'what we know about the mind or brain'. We will examine these arguments and 'what research shows' and glance at something of 'what we know about the mind or brain' below.

Collins refers in her discussion of the Key Competencies to the views of the psychologist, Gordon Stanley, but an examination of Stanley's position as presented in Collin's book reveals a significantly more measured rendition of the sceptical theory-based argument about cross-curricular competencies and the domain specificity of human capabilities than that offered by Collins (Stanley 1993).

Stanley sees 'research' as a crucial basis for his views and notes, quite rightly, that:

> The key publications from the Mayer Committee are remarkable for the absence of reference to the professional and technical literature on competency-based education (p.145).

Stanley accurately describes the change from the conceptions of the Finn Committee to that of Mayer as a change from 'achievements to abilities', and in doing so he references his comment to psychometrics and cognitive psychology (the literature of abilities) rather than the literature about competency-based education and training (p.147).

Under the section heading 'Specificity versus Generality of Ability', Stanley reviews the question of the 'breadth' of ability as 'one which has been of considerable interest to cognitive psychologists.' He summarises the situation by saying that:

> The traditional idea of general ability, which was unquestioned for much of this century, has more recently been decomposed into a range of broad generic abilities (Gardner 1983). For instance, the psychometric and neuropsychological evidence indicates that verbal and visual-spatial skills are not part of a more general ability (intelligence) but are independent of each other and need to he considered different, indeed orthogonal factors of ability. This means that performance in one domain is unconnected to ability in the other domain. Models of educational outcome which seek common benchmarks need to face the fact that achievement at some benchmark of competence in verbal skills may not always he achievable by someone who is able to exceed a 'comparable' benchmark performance in mechanical skills. Other generic abilities may be neuropsychologically independent in the same way.
>
> Often employers complain about the unevenness of educational achievement in individuals without recognising that this 'unevenness' may be more a reflection of inherent individual ability patterns than educational experience (p.148).

Stanley makes the unexceptionable claim that:

> People who excel in verbal skills may have limited achievements in certain practical visual-spatial skills and vice versa (p.148).

Unlike Collins, Stanley's summarising statement recognises that there has been no determination of the issues:

> Within the ability literature, efforts to specify basic cognitive processes that have generality have been inconclusive (p.148).

He summarises the evidence as meaning that 'caution should be exercised in attempts at generic specification of ability' and that:

> The literature from differential psychology implies that it may be harder to specify generic competencies in a reliable and valid way than envisaged by Mayer. Carroll (1992) points out that:
>
> > 'Establishing an ability, or a measure of that ability, is, in the first instance, a problem of determining what kind of tasks, and what attributes of those tasks, should be involved in its measurement. It is also a problem of determining whether the measure has certain characteristics, for example, high homogeneity and reliability, according to the psychometric model. ... As yet few of the traditionally recognized abilities have been subjected to this kind of analysis' (p.267).
>
> Carroll's conclusion reached, after many decades of ability research, is somewhat sobering to those who consider valid measurement can be easily attained for Mayer type competencies. (p.148)

Under the headings of 'Competence and Expertise' and 'Transfer of Learning', Stanley reviews some of the research on the differences between novices and experts and the importance of knowledge in expert performance:

> The expertise literature demonstrates clearly that there are domain specific skills which are not easily transferred across domains. Such skills are used by experts to provide efficient problem solutions. In the absence of specialised knowledge, non-domain specific skills are used involving what has been termed 'weak' problem solving strategies and include generate-and-test procedures, trial-and-error search, means-end analysis and problem reduction (Anzal 1991, p.65). From the perspective of this literature, these generalisable basic principles which transfer across situations are useful only because one does not know anything specific. They allow weaker solutions than those derived from expert knowledge (p. 151).

And he concludes that notions of general strategies have been 'oversold':

> The message from the literature on transfer of training is that the idea of general strategies or competencies has been oversold. They are no substitute for the building up of knowledge bases in specific domains. The evidence emerging from a number of recent cognitive analyses is even stronger. It suggests that ways of thinking applicable for one domain of knowledge may be inapplicable in another (p.151).
>
> Aspects of CBE find support in contemporary learning theory and cognitive science. Nevertheless such research confirms that the process of learning and the interrelationships between ability, instructional paradigm and knowledge domains are complex, as are outcomes. Such findings imply that we have not reached a stage where one model should he given priority status over others. The current literature on learning and instruction suggests that any attempt to impose a strong evaluative structure on the secondary or TAFE curriculum to give priority to a common set of generic competencies should be approached with caution. There appear to be good grounds, from existing cognition research, to be skeptical about the likely efficacy of the generic competencies enterprise in the form currently proposed by the Mayer report (p.152).

A Response to Stanley

The position presented by Stanley is the most coherent, the best informed and the best formed of the research-based arguments about notions of generic skills and the Key Competencies produced in Australia. He presents in some detail the reservations suggested by McGaw that came to shape the rejection of assessment of the Key Competencies in New South Wales. The views and arguments outlined by Stanley are related to the questions raised by the ACACA when considering the challenge of the Mayer proposals. As will be shown in Chapter 6, Stanley refers to a literature that is substantially at odds with his own line of argument, and it should be noted that his conclusions are, in any case, no more than measured reservations that he claims should prompt caution. They certainly do not support the claim of Collins that 'empirical evidence' has shown the Key Competencies are 'hocus pocus'.

In the review of both the international Key Competencies movement in Chapter 2 and the criticisms of Key Competencies in this chapter it is contended that:

- generic skills are not competencies;
- nor are they appropriately thought of as skills but they are appropriately seen as generic abilities;
- the claim that research has demonstrated that Key Competencies are invalid or illusory is not strong;
- there is a lack of rigor in claims that objections to the Key Competencies are based on 'what we know about the brain'; and
- some of the theoretical and research based argument against the Key Competencies does not show the conceptual rigor that is the special province and responsibility of theoreticians and researchers.

It is appropriate to look rather more widely than Stanley has at the exploration of abilities undertaken in the literature of psychometrics and cognitive psychology to which he partially refers. Before examining this theory and research into cognitive abilities, it would be appropriate to explore and try to clarify some basic conceptual issues.

4
The Concepts of Ability and Competence

What is a Cognitive Ability?

Chapters 2 and 3 have reviewed the development of the generic and work-related skills movement in Australia and internationally. Chapter 3 outlines how advocates promoted the value of the Mayer Key Competencies and critics have challenged them, in particular how critics challenged the Key Competencies as at odds with 'theory' and 'research'. Chapter 6 will review the theoretical and empirical exploration of human abilities in psychology, the field that has systematically explored generic and transdomainal abilities over the last 110 or so years. This body of theory and empirical research seemed to be invoked by Collins in dismissing the Key Competencies, and it was explicitly invoked by Stanley in making his more cautiously qualified challenges to the Key Competencies. This body of theory and research is also effectively rejected as a basis for developing key competencies by Weinert and DeSeCo. The position of Weiner will be examined in the next chapter, and what psychometrics and cognitive psychology suggest about notions like the Key Competencies will be examined in chapter 6.

Before reviewing this theory and research, it would be appropriate to deal with some conceptual issues and to give some attention to what we might mean by a cognitive ability and to different approaches to envisaging and understanding such abilities. It would also be appropriate to examine the relationships between notions of ability and notions of key or general competencies. The different associations that surround a word like 'competence' and the notion of 'competencies', for instance, have an impact on the coherence of educational discourse and of educational policy.

Similarly, the attribution of kinds of ability is fundamental to all assessments in education and psychology, but the notion of an ability or a set of abilities is elusive and ambiguous. Such is the fate of many of the most pervasive terms (the term personality is an even more egregious example), and such equivocalness of terminology can be the basis of much confusion. Attempting to clarify and rationalise such key terms is, or should be, a crucial task for theoreticians and researchers.

What do we mean when we talk about an ability or attribute an ability to a person?

What do we understand, for instance, by the word 'ability' in the following statements?

- She has the ability to do such and such.
- He has the ability but not the inclination to do such and such.
- The ability to do such and such is fundamental to such and such.
- It takes a special ability to do such and such.

What, if anything, is this ability? What, if anything, does this term 'ability' refer to?

While these questions may seem of merely philosophic interest, different answers to them actually have a significant impact on what counts as significant and convincing evidence for an ability; and whether and in what sense an ability can be defined operationally and thus, 'measured'.

Competence and Competencies

While the nature and definition of competence and competencies have been matters of debate, it can be argued that the notions of competence and competencies are not particularly difficult. Competence is a hypothesised level of performance, and a competency is an area of performance in which the level of competence must be reached. These issues can be thus clarified and resolved as matters of definition.

As competence is usually understood in terms of a particular domain, defining areas of competence is usually not particularly problematic. Stating the areas in which a nurse or a plumber needs to be competent is not difficult. The really formidable difficulty for competency-based assessments is in operationalising the distinction between competence and incompetence in different areas. To stipulate such a difference is not, of course, to establish or demonstrate it.

However, the Key Competencies (in spite of their title) are generic abilities rather than particular knowledge domains or classes of tasks. They may be thought of as the broadest of areas of competence but are not, as we have seen in the discussion of the Mayer levels in Chapter 2, anything like a serious definition of competence. In exploring the Key Competencies we are examining generic skills or general abilities, and in doing so we have to adduce a theory of general abilities. Such a theory will be an account of the underpinnings of performance and it will also involve an account of learning. In other words, if we are to understand the Key Competencies and other concepts of generic abilities, we have to

analyse what we mean by the word ability and what we mean when we attribute general and underpinning abilities.

The Term Ability

A number of different possible meanings or assumptions can be discerned behind the attribution of an ability to an individual. The noun ability is defined in the following ways in the Macquarie Dictionary.

1 the power or capacity to do or act in any relation.

2 competence in any occupation or field of action, from the possession of capacity, skill, means or other qualifications.

3 talents, mental gifts or endowments.

These definitions focus on a power, capacity, skill, talent, gift or endowment of an individual, and they seem to imply that this power or capacity is a kind of system or mechanism possessed by, or part of, the individual. Similarly, in common parlance, an ability is understood as a capacity of the individual, that is the power an individual has or can generate to undertake performances. On the other hand, the way in which the concept of ability has been understood and defined within the scientific discourse of psychology has been a matter of contention, and these contentions have involved some important differences of methodology and outcome.

How might we further analyse this capacity or capacities?

By way of introduction to these issues, it is useful to distinguish four different kinds of response to the question about the nature of ability that view it as either:

- a mental mechanism or capacity that underpins performance;
- a characteristic of a domain or class of tasks or kinds of observable performance;
- the knowledge and understanding of a domain; or
- a notion or construct that organises our ideas or expresses our intentions.

Each of these notions will be examined below. The different perspectives are not mutually exclusive, and a comprehensive model of performance, in our current state of knowledge, will account for and integrate different approaches to and understandings of ability.

Different Approaches to Defining Ability

The first of these perspectives on ability, which might be termed the cognitivist position, follows the dictionary definition and is focused on brain functions and capacities. The nature of these mind/brain capacities is uncertain in our current state of theory and empirical research. Although the psycho-biology of the mind/brain is a matter of much research, links between higher cognitive processes and basic cognitive processes (let alone biological processes) have not been established (Greenfield 1997; Calvin 2001; OECD 2002). The cognitivist position on abilities has been subject to the behaviourist challenge that the mentalist notions of a cognitivist position are both unfounded and unnecessary. The key tenet of such behaviourism is that 'the only appropriate subject matter for scientific psychological investigation is observable, measurable behaviour' (Reber 1985).

A second approach to ability views it as a characteristic of a class of tasks. This approach focuses on kinds of observable performance and the ability to do something, and hence it might be called a behaviourist, or crypto-behaviourist approach to defining ability. This crypto-behaviourist approach seeks to side step the problematic issue of the capacity that enables someone to do something (and the awkward issue of the different ways individuals may do the same thing) and focuses on the apparently observable characteristics of what is done. So this behaviourist approach defines ability in terms of the characteristic of a class of tasks or kinds of observable performance.

A third way of understanding ability sees it as the knowledge and skills of a particular domain. A domain in psychological parlance means an area or, as defined in the Macquarie Dictionary, 'the scope or range of any sphere of personal knowledge'. A domain-related concept of ability emphasises the place of knowledge in performance, in contrast with trans-domainal notions of a capacity-based understanding of ability.

The differences between a crypto-behaviourist, a cognitivist and a domain-related approach to the notion of ability evolved, and has played out a dialectical struggle, during the scientific exploration of ability in the C20 (Gardner 1987; Hunt 1993). Chapter 6 will examine how the exploration of abilities was dominated by the psychometric approach for much of the C20.

The Empiricism and Operationalism of Psychometrics

Psychometrics is the branch of psychology (or applied psychology) that is concerned with 'the application of the principles of mathematics and statistics to the data of psychology' (Reber 1985). As its name suggests, psychometrics is concerned with what is called

'measurement', and hence it claims to be empirical rather than rational and theoretical. In its claim to empiricism, as distinct from 'arm chair theory' (Cattell 1987) psychometrics tends towards operationalism in that it is based on 'procedures carried out without regard to any theory' (Reber 1985). One can pursue psychometric research without any particular theory about the notion under examination, and this atheoretical aspect of psychometrics was cognate with the hardline behaviourism which took an avowedly operationalist stance. Operationalism (related as it is to positivism) argued that 'the concepts of science be operationalised – that they be defined by, and their meaning limited to, the concrete operations used in their measurement' (Reber 1985). The behaviourist repudiation of mental phenomenon was consonant with the psychometric approach to understanding cognitive performance that is based on the statistical analysis of test performance. Behaviourism and psychometrics were interrelated in their genesis and growth.

In bald outline, the development of a body of theory and evidence within psychology about cognitive abilities reflects the predominance of a psychometric approach for the first 50 or 60 years of the C20. Both behaviourism and the psychometric approach were subject to the cognitivist challenge during what came to be known as 'the cognitive revolution' of the 1960s and 1970s (Gardner 1987). Psychometrics, in particular, was challenged by cognitivists for its refusal or inability to explain cognitive processes. The psychometric view of abilities was also challenged for its lack of recognition of the fundamental importance of knowledge in performance.

A second wave of criticism of psychometrics can be discerned in the 1970s and 1980s. This criticism challenged the political nature and the social consequences of the psychometric approach to understanding ability. This socio-political challenge to psychometrics not only drew attention to the limits of the psychometric approach to ability, it also challenged the soundness of an approach based merely on the statistical analysis of test performance. Some criticisms went further by saying that not only were psychometric methods limited, they were fundamentally fallacious in that they claimed or implied they established the structure of the mind/brain (Gould 1981). It was charged by some critics that this claim or suggestion about the structure of abilities based on psychometric evidence amounted to the reification of statistical artefacts as cognitive structures.

These criticisms of the psychometric approach are telling, but, at its most careful and most rigorous, psychometrics chooses to avoid accusations of reification by defining ability strictly in terms of the characteristics of a class of tasks, the second aspect of an ability described above. As with other behaviourist approaches, this version of the psychometric approach claims to be firmly founded on observable behaviour, and hence it is apparently

concrete. And this circumscribed psychometric approach does not make any claims about cognitive processes and structures.

The cognitivist approach, on the other hand, might stigmatise the definition of an ability as a class of tasks as an evasive 'black box' theory (Reber 1985; Bullock, Stallybrass et al. 1988). It might criticise such behaviourist definitions as lacking explanatory power because they treat the mind/brain as an uninterpretable black box (Sternberg 1977). These critics argue for the importance of exploring and explaining cognitive processes. They might reject the metaphysical notion of a class of tasks and aim for a psycho-biological explanation of ability (Gardner 1993). That is, they might aim to explain the cognitive capacities that account for performance across different kinds and classes of tasks.

The difference between an ability as a class of tasks or as a faculty/capacity reflects the differences between behaviourism and cognitive psychology. The behaviourist and cognitivist approaches to notions of ability can be related to our usual assumptions about learning and performance so that it can be seen that a theory of ability is also a kind of theory of learning and performance.

Learning, Performance and Ability

To learn is primarily defined in the Macquarie Dictionary as: 'to acquire knowledge of or skill in by study, instruction or experience.' What is commonly assumed about the basis and nature of learning? What is commonly assumed about the basis of the performances that result from learning?

We generally assume that learning is the knowledge and skills absorbed/gathered/ processed/stored from experience. We assume that performance is a result of prior experience in an activity and the extent to which an individual has absorbed/gathered/processed/stored previous experience. We assume that individuals differ in their capacity, aptitude or ability to learn, and we also assume that individual's capacity, aptitude or ability to learn is different for different kinds of things. That is, equal inputs in experience for different individuals will not result necessarily in equal outcomes in learning. There will be an interaction between the biological capacities and the temperament of the individual and their experience that will be manifest in their ability to perform.

Kinds of Knowledge

It is usual to think of learning as consisting of knowledge and skill. This conceptualisation seems to presume that knowledge is some store of data derived largely, although not necessarily or exclusively, from experience. Skill, on the other hand, while it involves and depends on knowledge, is conceptualised as the ability to do something. This distinction between knowledge and skill means that it is possible to describe someone as knowing something, and even knowing how to do something, but as not being able to do that something. Alternatively, it is also the case that it is possible to be able to do something and not to know, propositionally, how to do it. The potential inconsistency here (someone knowing how to do something but not being able to do it) can be avoided by conceptualising all learning as a kind of knowledge and hypothesising two kinds of knowledge as underpinning all performance. According to this conceptualisation there is declaratory knowledge (knowing that and what) and procedural knowledge (knowing how) (Gagne, Yekovich et al. 2002).

An alternate notion conceptualises performance as a matter of deploying cognitive capacity. This notion of performance commonly involves a distinction between crystallised abilities (a store of knowledge and experiences) and fluid abilities (capacities used to deal with new or unfamiliar problems) (Cattell 1987; Horn 1994). This second view has gained increasing prominence in psychometrics and cognitive psychology in the past 30 years, and it is the basis of a good deal of current thinking about abilities. To such a view, abilities involve knowledge but they are also underpinned by flexible capacities that are applied to different kinds of tasks. This view sees these fluid abilities as transdomainal (and in that sense generic), or as being the flexible, general capacities that underpin learning and performance.

The Specific Knowledge and the Generic Skills Views

On one hand we have the contextualised learning and skills view that everything is more or less the specific knowledge and skills of a kind of activity. According to this view there is little that is general about performance and there is little transference from one context of learning or particular skill to another. On the other hand there is the generic skills view of learning which takes the position that while specific knowledge (crystallised ability) is important to performance, performance (particularly unfamiliar or novel performance) calls on general skills (fluid abilities?) of the individual that are turned into specific abilities (crystallised abilities). According to this view, the same cognitive mechanisms are used in most kinds of thinking, learning and performance.

It is possible to conceptualise/theorise all abilities as more or less specific. Such theory entails little or no generalisation across tasks or domains. (See the arguments of Gardner and Ceci in Chapter 6.) Such an approach to abilities is not concerned to, and does not, generalise. On the other hand, envisaging performance as entailing crystallised and fluid capacities posits generalised or transdomainal abilities. (See Cattell, Sternberg and Carroll in Chapter 6.) Although there are sharp contentions against generic or transdomainal skills, the notion of generic skills (particularly as represented in the notions of crystallised and fluid abilities) is the predominant metaphor used to understand abilities in psychometrics and psychology. These psychometric approaches to ability are (or claim to be, while anyone is watching) resolutely agnostic about capacities or mechanisms and remain carefully focused (in a behaviourist fashion) on tasks.

With these different considerations in mind, let us return to the question about the nature of ability and the different responses to this issue.

The Class of Tasks View

It can be argued that at its most literal an 'ability' is the ability to do something, and hence it might be contended that at its most basic the attribution of ability is a claim about the something done. So for this definition an ability is a kind of thing that has been done.

This approach can be strictly circumscribed so as to avoid the difficult issue of how something is done and hence the charge of reification of the hypothetical as real. It also conveniently avoids the difficult issue of how things are done by treating the mind/brain (in a behaviourist fashion) as an opaque black box, the workings of which are not, and need not be understood.

In this crypto-behaviourist sense, the attribution of an ability does not necessarily claim anything in particular about how the thing is done by an individual. It claims rather that the thing done is a kind of thing or a class of performances or tasks. Hence this might be called a class of tasks-based definition of ability. The challenge for this approach is to offer a definition or basis for determining class of tasks.

A black box behaviourist stance is most intelligible and satisfactory where the class of tasks is very constrained. That is, where the task poses a problem with a unique solution that can only be arrived at by a unique and specific process of reasoning. The classic fluid intelligence tasks like Raven's matrices, for instance, are just such constrained tasks. (See appendix 3 for an example of Raven's Progressive Matrices.) Even the classic verbal

analogy questions (which are apparently more matters of crystallised intelligence than Raven's matrices) are processing within unique (and arbitrary) constraints.

However, much thinking (perhaps most thinking?) is not like the closed and constrained logical processing involved in Raven's matrices and verbal analogies. (See Sternberg's view of real world problems outlined in Chapter 6.) Leaving aside the fact that much thinking is not a matter of solving a particular problem (and certainly not a problem with a constrained process for deriving a unique solution), many problems can be solved in several different ways. These different ways of solving a problem can involve different abilities, and different individuals can use different abilities in solving the same problem. Hence the relationship between capacities and classes of tasks itself becomes problematic.

What is a Class of Tasks?

It is evident that there is a very, very broad range of possible task sets (entailing a very broad range of knowledge and skills), but it is not evident (and it is not even likely?) that the performances of individuals are somehow specific to particular tasks. There is reason to ask whether the capacities or faculties of individuals (that is, abilities defined as capacities of individuals) are more or less specific to tasks or are in some sense expressions of general capacities or faculties applied to different kinds of tasks involving different stores of knowledge. So tracing through the implications of the class of tasks definition of ability ends up enmeshed in just those questions about the capacities of the individual that this approach sought to avoid. Nevertheless, such a class of tasks based definition of abilities has been commonly accepted in the psychometric tradition.

For instance, Carroll, in one of the key texts to be discussed in the next chapter (Carroll 1993), gives a classic statement of the class of tasks position when he defines ability in the following terms:

> As used to define an attribute of individuals, ability refers to the possible variations over individuals in the liminal levels of task difficulty (or in derived measurements based on such liminal levels) at which, on any given occasion in which all conditions appear favourable, individuals perform successfully on a defined class of tasks (p.8).

Leaving the gesture at ability as an attribute of individuals, this is a definition of ability in terms of a 'defined class of tasks'. This definition seeks to avoid the problems of defining the way in which someone does something by focusing on the apparent objectivity of what is done, but such a definition only leads to the question about the basis for defining a 'class of tasks'.

- What is a class of tasks?
- What is the basis for determining the membership of the class?
- How do we determine that these tasks are a class distinguished from other classes?

Carroll goes on to define a class of tasks in the following terms:

> By class of tasks, we mean a group or series of possible tasks that have at least some identical or similar attributes. These attributes may refer to the kinds of stimuli that can be dealt with, the kinds of action that must be performed, or the means by which those actions can be performed (p.8).

Thus, being a class means that tasks are in some way identical or similar, and this is explained as having some similarity in kinds of stimulus, kind of action or means by which actions are performed. But the criterion of having 'at least some identical or similar attributes' in Carroll's definition is rather impressionistic. So it seems that the supposedly evident and observable comes down to a rather vague attribute.

This class of task definition of abilities is carefully circumscribed in that it only attributes (directly at least) characteristics to tasks. It does not claim to attribute characteristics or capacities to all test takers or to any individual test taker. This characteristic caution on the part of the class of task definitions of ability allows those who take this stance to repudiate charges of reifying statistical artefacts as psycho-biological facts about groups or individuals (Carroll 1995).

In Chapter 6 we will see that the charge of reification of statistical artefacts is a fundamental challenge made to the psychometric approach to abilities. Carroll would contend that his definition of an ability as a class of tasks does not fall into the trap of reification. Some psychometric practitioners are not always as cautious as Carroll, and a position commonly held by some psychometricians, known as latent trait theory, is such a reification of statistical artefacts. According to latent trait theory, test items are 'observations' that may, if they have certain statistical characteristics, coherently define a variable. But what is this variable? In dealing with this issue, latent trait theory takes the step that Carroll avoids, because it attributes a trait to the testees rather than making a claim about the test items. Carroll will attribute the ability to do the items to the candidates, but the ability is defined in terms of the characteristics of the items. Hence Carroll can claim that he is not reifying a statistical artefact.

The advocates of latent trait theory are not so circumspect (Anastasi 1988). According to latent trait theorists, an ability is an inferred characteristic of the testees, and hence they are open to the charge of reifying statistical artefacts. According to latent trait theory, the

variable of aggregated performance on test items is a varying trait of the candidates, and at its most extreme, a 'measurable' quantity (Kline 1991; Jensen 1998; Kline 1998).

While the class of tasks approach of Carroll is fairly carefully circumscribed, and can reject the charge of reification, it is also limited. The class of tasks approach can make no claim about the way different individuals perform a task, that is about the capacities of the individual. The class of tasks approach to defining abilities is behaviourist in that it claims to be a matter of observable behaviour, and the mechanisms of performance remain an unexplained 'black box'. This kind of psychometric approach seeks to side step questions of definition and instead makes operational definitions on the basis of statistical coherence. Such an approach has been increasingly subject to challenge as atheoretical or merely operational (Sternberg 1991), and more complete accounts of performance have been demanded by critics of the psychometric approach. Thus, the psychometric approach was challenged to explain the way individuals performed tasks; and how these performances are characteristics of individuals. It was challenged to explain what individuals did in completing a task and to explain what mechanisms enable individuals to complete tasks (Sternberg 1977; Sternberg 1985; Sternberg 1991).

A Cognitive and Domain-related View of Ability

The apparent tightness of a class of tasks view of ability is vulnerable to the charge of narrowness and the claim that fundamental issues are not addressed by such an account. Such crypto-behaviourist views of ability were the basis of the psychometric approach, and remain the defense against charges of reification of mathematical factors currently made by critics of the psychometric and factor analytic approach to defining abilities.

The 1950s and 1960s saw increasing challenges to the behaviourist orthodoxy in psychology in the English speaking world. Treating the mind/brain as a black box no longer seemed satisfactory, and researchers and theorists wanted to know how the mind/brain processed ideas and solved problems. This cognitive revolution, as it came to be known, gave rise to research into cognitive processes and into faculty or modular theories of mind/brain (Sternberg 1985).

These cognitivist views sought to define abilities (and in Gardner's case 'intelligences') in terms of cognitive processes or mental modules, faculties or mechanisms (Fodor 1983). These approaches sought to explain performance and to define ability in terms of cognitive processes or capacities. They were rejections of the strictures of behaviourism and they were also in some cases

rejections of the psychometric approach to defining abilities. Those taking a cognitive approach criticised the opaque, agnostic limitation of the class of tasks approach to defining abilities, and criticised the more naïve latent trait theories as reification. At their most optimistic these cognitive approaches hoped to describe and explain thought processes and to work towards a meeting and synthesis of higher psychological functions with psycho-biology (Hunt 1993).

This reaction of cognitive psychology against behaviourism and psychometrics will be sketched in Chapter 6. For the moment, it should be recognised as a different view and definition of ability from that offered by the behaviourist class of task definitions or latent trait theory of ability, common in psychometrics.

As well as the cognitivist challenge to psychometric views of ability, a knowledge-based or domain-related view of ability developed in the 1980s. This view argued that performance was more or less dependent on the knowledge and experience in a domain. The domain-related view challenged both the class of tasks and the cognitive capacities views of ability.

A Constructivist View of Ability

There is one way of avoiding the charges of atheoretical narrowness or reification to which the class of tasks and the latent trait definitions of ability are vulnerable, without taking on the very challenging task of locating abilities in mind/brain functions, processes, capacities or modules. Under the influence of certain philosophic and sociological approaches, this position views abilities as notions, concepts or constructs (Flew 1983; Bullock, Stallybrass et al. 1988). Viewing abilities as constructs avoids the claim of reification as no claim is made as to the physical existence of processes or abilities. This construct notion of an ability implies that performance is in some undetermined (and perhaps indeterminate) interaction of tasks, capacities of individuals and domains. It does not make any empirical claims about the inherent nature of the tasks (a metaphysical claim about a class of tasks) nor make any particular claim about the capacities or mental mechanisms of the testees. But this constructivist view of abilities is prone to the challenge about a basis for the definition of an ability. How then can a constructivist view of ability be developed? What might be the constructivist criteria for an ability?

Two potentially supporting criteria might be offered for such a constructivist definition of an ability. It might be claimed that this constructivist ability would need to be logically coherent and pragmatically useful. The logical coherence of the construct would be a matter of the presentation of a clear conceptual framework of meaningful distinctions. The pragmatic usefulness of a construct might depend (in the manner of the

psychometric approach) on meaningfully discriminating among and within the performances of individuals, and that these discriminations would need to be reliable and have predictive validity. Basing a constructivist approach on these positions unites it with the psychometric approach and the empirical grounding claimed by the psychometric approach. Such a constructivist approach contrasts with the class of tasks approach in that it does not claim that its hypothesised kinds of ability are any more objective that its hypothesised kinds of task.

The grand purpose of cognitive science is the integration of psycho-biology with psychology, so that basic cognitive processes are integrated with higher processes (OECD 2002). This integration has not yet been made, and in its absence, a constructivist approach drawing on and integrating the theories of cognitive process and the pragmatism of psychometrics seems the best stance and is perhaps the best way forward in understanding kinds of ability. However, such a constructivist reconciliation is a weak position in that it does not meet the challenge of attaining a psycho-biological view of cognition.

Fact and Artifact, Theory and Empiricism, Reification and Reality

The claims for and against the class of tasks, the capacity or faculty, the domain-related and the constructivist approaches to defining abilities involve some of the fundamental methodological distinctions to be made in contemporary theory and research. The class of tasks approach is strong if it is circumscribed to be only a class of tasks, but how and when something becomes a class of tasks is a matter of judgement and is always open to dispute. The latent trait approach is subject to the criticism that it reifies a class of tasks into a capacity or faculty. On the other hand it might be argued that the class of tasks and the latent trait theories of the psychometric approach are empirical, that factors or item response theory measurement scales are facts and are in that sense 'real' (Jensen 1998). It seems that some advocates of the psychometric approach on one hand deny that they reify factors or measurement scales, but, on the other hand, they are inclined to view their approach as empirical and factors and measurement scales as 'real'.

Such claims can be challenged on the grounds that the results of factor analysis and item response scaling are statistical artefacts rather than facts, that they are mathematical fabrications and are not real, and that the construction of such statistical artefacts does not amount to an empirical concern with fact. These charges amount to the claim that 'measurement' in psychology and education is metaphorical rather than the literal measurement of physics. On the other hand, capacity or faculty definitions of ability can be said to be empirical in their concern with physical fact, but the reality of separate mental

capacities and faculties is elusive and currently have the status of hypotheses that can be neither proved nor disproved.

An example might elucidate these differences. Critics of Gardner's multiple intelligences theory from within the psychometric tradition describe his theory as lacking empirical grounding because it is not open to or has not been offered empirical proof or refutation (G.Kearsley 1994; Carson 2001). Gardner denies these charges on the basis that his theory is grounded in research about brain functions and systems (among other things) and that the theory offers hypotheses about how the brain is organised and functions. For his part, Gardner sees the psychometric approach as lacking empirical grounding in the mind/brain, and he sees his own theory as offering a view of mind/brain structure and functioning that has the potential to be proved or disproved (Gardner 1993).

A constructivist approach contrasts with the empiricism of the capacity or faculty approach in that it does not claim its notions are real. It claims rather that they are useful, and there can be a pragmatic synthesis of the psychometric and the constructivist views. A scientific theory of ability, on the other hand, would entail claims about the performance of individuals that are based on a view of the way the mind/brain works and, possibly, the way mind/brain processes relate to different kinds of tasks.

Abilities and Competencies

If we look back from this discussion to the notion of competencies it is evident that competencies may be understood as domain-specific knowledge and skills, but there are many possible domains and such a conceptualization offers little or nothing that is generic. It seems that the notion of generic competencies is best understood as classes of tasks, that is classes of things that one needs to be competent in. Class of task based definitions of competencies are prone to challenge on the basis of the criteria for determination of a class of tasks and their inability to account for how individuals perform classes of tasks. Class of task definitions and knowledge-based definitions of competencies are also vulnerable to the claim that certain general underlying capacities account for performance in what may appear to be different classes of tasks and domains. Being a salesperson and a psychological therapist are different kinds of activity, for instance, but both kinds of tasks are based on, in Gardner's terms, interpersonal intelligence.

Competencies are usually defined as classes of tasks, and the Key Competencies, for instance, aspire to be kinds of tasks. The Key Competencies are not abilities in that they are not notions of cognitive capacity, and the Key Competencies do not draw on any theory of

ability or theory of classes of tasks. Yet, a set of Key Competencies (that is a set of generic skills) is a kind of ability theory, and should be understood and developed as such.

The development of the Key Competencies was based on a question about what people needed to be able to do, and hence they tend to be kinds of activity or classes of task. The research of the Mayer Committee does not appear to have made contact with or been influenced by any ability theory, and the Key Competencies do not draw on what is known about cognitive processes and systems. Thus the Key Competencies are atheoretical.

With this criticism in mind it is worth considering what might be involved in a theoretically sound and scientifically-based theory of ability.

Towards A Differential Ability Theory

An ability theory aims to give a general explanation of cognitive performance and to discern the major kinds of ability to perform. A comprehensive ability theory will deal with:

- the ways social context shapes performance;
- the nature and demands of different kinds of performance; and
- the biological basis of performance.

Such a comprehensive differential ability theory would aim to define a small number of different abilities that covered as broad a range of performances as possible with the smallest amount of overlap between different abilities. A differential ability theory would thus aim to be both comprehensive and parsimonious.

A psychometric model of ability would be based on similarities and differences in performance between individuals and the extent to which the hypothesised ability/abilities can explain the variance between individuals in a set of performances on classes of tasks. A psychometric model depends on patterns of covariance in test performance, and psychometric claims about a general or a group of specific abilities are based on statistical measures of internal consistency in test performance. Whether, or the extent to which, psychometric theories of ability are empirically based or scientific is a matter of dispute.

Epistemological and Curriculum Theories

The approaches to defining abilities discussed above and the theories outlined in Chapter 6 are drawn from the field of psychology where there has been an attempt to develop a theory of human abilities. Some other areas of intellectual endeavour have sought to develop

theories of knowledge that are also in some sense theories of ability, although they do not have the general and empirical focus of the theorising and research in psychology.

The segment of philosophy knows as epistemology has concerned itself with the nature of knowledge and the basis of knowledge. This area of philosophic conceptualisation might seem to offer a categorisation of knowledge or fields of knowledge, but it remains purely conceptual or logical and is without empirical grounding in actual human performance or mind/brain processes. An example of such work can be seen in philosophically oriented theorists of curriculum such as Phenix (Phenix 1964), or even in Bloom's taxonomy (Bloom, Krathwohl et al. 1956).

Educational theory has also concerned itself with the delineation of kinds of knowledge through curriculum theories, but curriculum areas are specific domains of knowledge or activity rather than the kinds of generic abilities or cognitive capacities discussed above. A body of knowledge or a domain is not an ability in the sense of being a part of a theory of ability or generic capacities. A theory of ability is a theory of performance rather than a theory of knowledge. A basic feature of a theory of ability is a concern with that which can be distinguished from knowledge which underpins performance. A theory of ability in this sense is concerned with aptitude ('capacity' in terms of this discussion) to learn, rather than with what has been learned or achieved. Just as focusing on aptitude (as distinct from knowledge-based achievements) involves minimising the importance of prior knowledge, so a theory of abilities involves a focus on those elements of performance that are in some sense separable from knowledge-based achievements. These distinctions are represented in the following diagram of what might be called a model or theory of learning and performance.

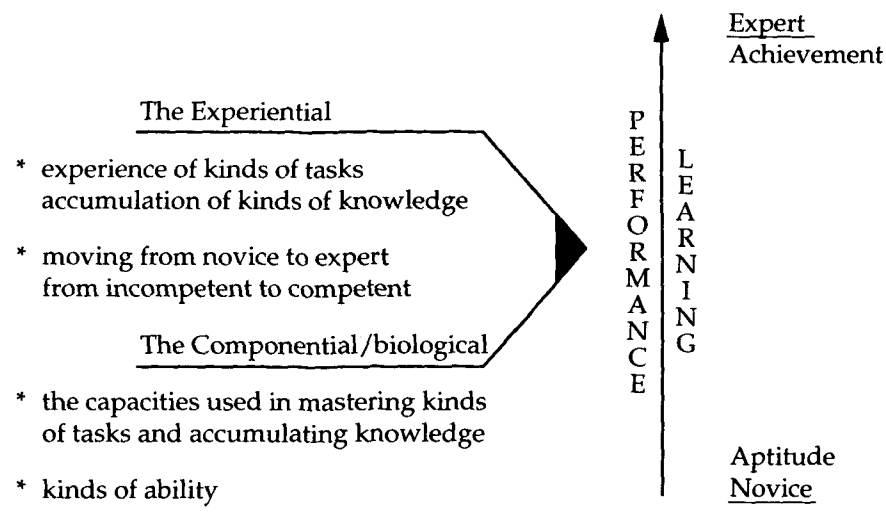


Figure 3 A Model of Performance (After Sternberg's Triarchic Theory)

This model seeks to explain performance, and it does so by the two horizontal lines on the left that converge in an arrow head as 'performance'. The top line comes from what Sternberg calls the experiential sub-theory, and it involves kinds of task and domains of knowledge (Sternberg 1985). In terms of Cattell's fluid/crystallised distinction, this is the area of crystallised abilities, and these abilities are understood as either specific or generic. The specific abilities are thought of as more or less propositional knowledge, and the generic are thought of as non-propositional skills. Also accounting for performance is the lower line which deals with the componential sub-theory of Sternberg and the biological aspect of Ceci's bio-ecological theory (Ceci 1990). This complex of ideas is concerned with abilities as cognitive capacities, and it involves fluid abilities and generic cognitive capacities.

This view of performance is in turn related to the ascending arrow of learning, and learning is related to the notion of a novice at the base and an expert at the top. The novice is paired with aptitude and reflects the componential and biological, whereas the expert is paired with achievement and reflects the experiential.

Although we readily think of curriculum areas or other areas of specialisation in education as entailing kinds of abilities (and we sometimes think of them as more or less special abilities), they do not constitute abilities in the sense of an ability theory outlined above. How then might we distinguish the kinds of ability (as distinct from the body of knowledge) developed in geography, biology and economics? While curriculum areas focus on different domains concerned with different ideas or phenomena, they do not as a result necessarily entail special or unique abilities to understand, reason about and solve problems in each particular domain. As in the relational model of Bowden and Masters (Bowden and Masters 1993)(and we will see similar ideas presented by others below), some contend that purportedly generic skills are actually domain or context specific and cannot be assessed separately from the domain or context. Undoubtedly, different curriculum areas call on and develop different kinds of ability and different kinds of ability are developed within an area at different points, but different subjects or curriculum areas are not distinct abilities in the sense outlined above. They are not general theories of cognitive capacities or even classes of tasks.

One might attempt to cluster, or to look for overlap, between the predominant curriculum-based abilities involved in different subjects or curriculum areas, but the subject areas themselves do not offer meaningful patterns of ability in themselves. A general theory of differential ability would have to be used or be formed from such a synthesis because the curriculum, or a curriculum theory, is not a theory of ability.

The difference between a curriculum theory and a theory of ability might be seen in a local and contemporary example of the curriculum framework based on the national goals of schooling offered by Piper (Piper 1997).

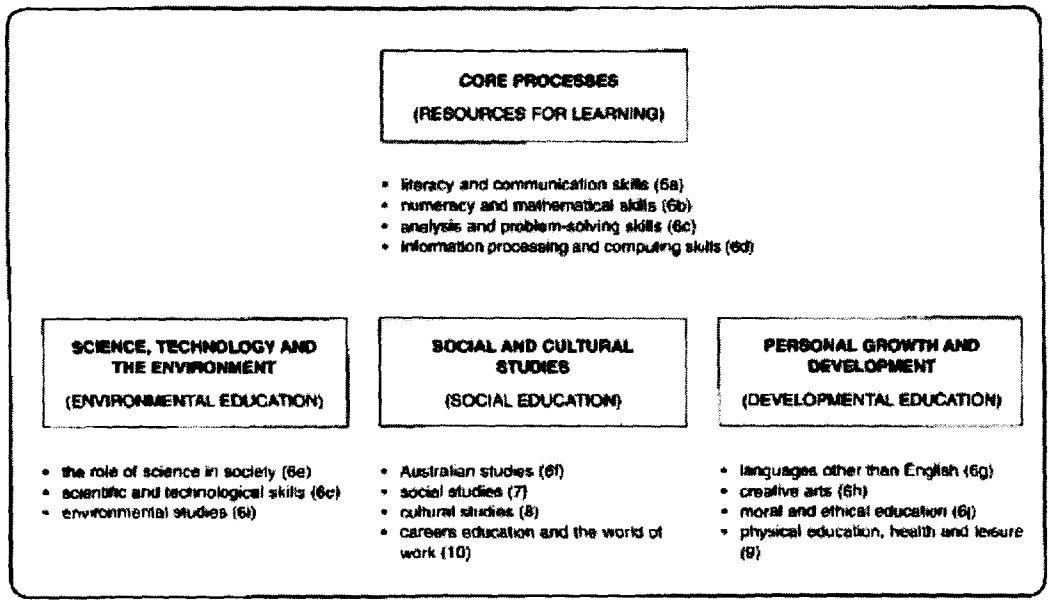


Figure 4 Piper's Skeletal Framework Based on the National Goals of Schooling

The centre of this common sense conception are the two broad content domains of Science, Technology and the Environment and Social and Cultural Studies. Cutting across these content domains are the core processes of literacy and numeracy. Problem solving and information processing and computing seem to be offered as equivalents of literacy and numeracy.

Literacy is evidently a generic and transdomainal complex of knowledge and skill. Literacy is a generic skill in the sense that it can be developed and exercised in all sorts of specific domains, and literacy skills can be applied to all sorts of specific domains. Numeracy can be thought of as a body of knowledge and skill, but it is not evidently generic and transdomainal in the sense that literacy is.

However, one of these core processes in Piper's model is of a different order altogether. The notion of a core process of analysis and problem solving is not readily or usefully thought of as an area of content or a body of knowledge and skills. (There is no subject on the school timetable called Analysis and Problem Solving.) The terms analysis and problem solving stand for generic thinking skills that are found in all subjects and would seem to be assumed to generalise across the curriculum. These thinking skills are not differentiated in terms of content or domain.

In contrast with both philosophic and educational theories, psychology has not focused on domains of knowledge or skill. Rather, psychology has focused on general or transdomainal skills that underpin performance, with a particular emphasis on the psycho-biological capacities or mechanisms that are the basis of performance.

Curriculum or Domain-specific Views of Ability

While is seems readily apparent that all sorts of assessment involves tacit views of cognitive abilities or capacities, most educational assessment is circumscribed by a domain, usually thought of as a subject or skill area. One might go further and argue that most educational assessment lacks a sense of particular cognitive abilities that together would add up to a comprehensive differential ability theory.

Accounts of what is assessed in different subject areas can seem rather simplistic. The members of a field commonly refer to the particular knowledge and skills of their domain as abilities, but they are clearly not theories of generic abilities in the sense at issue here. Hence teachers commonly talk about physics ability, literature ability or accounting ability. As Carroll bluntly states it, from the perspective of cognitive psychology and ability theory, these notions have no 'scientific meaning':

> 'Mathematical ability,' therefore, must be regarded as an inexact, unanalyzed popular concept that has no scientific meaning unless it is referred to the structure of abilities that compose it. It cannot be expected to constitute a higher-level ability. Similar statements could be made about other everyday concepts like 'musical ability,' 'artistic ability,' 'creative ability,' 'problem-solving ability,' and the like. This rigor is applied to the popular notion of intelligence also. For that matter, a similar conclusion can be reached regarding the concept of 'intelligence' which is also an inexact, unanalysed popular concept that has no scientific status unless it is stated to refer to the abilities that compose it, as described in the present volume (p.267).

Discussion of performance within particular domains does not attempt to generalise about different capacities. When an ability is identified in discussion within a domain, it is not understood to be general. Within domains there may be theories of performance, but they are not theories of ability (or more importantly differential ability theories) because they are not generalising across domains. A theory of ability as understood in this thesis on the other hand is a generalisation across domains.

There has been vigorous debate about the usefulness and validity of generalisations about abilities across domains for the whole history of cognitive psychology, and it can be seen in the debate about the Key Competencies that different views are part of the politics of educational assessment.

There is one exception to the paucity of coherent theory about ability in most educational discourse, and that is the ubiquitous Bloom's Taxonomy (Bloom, Krathwohl et al. 1956). This taxonomy has a very wide currency in educational thinking, but it is an analytic theory, in essence, a philosophical position rather than a scientific or empirical theory that seeks grounding in actual performance, let alone the capacities that underpin performance. One would imagine that Carroll, if asked, would describe Bloom's taxonomy as another set of unanalysed popular concepts that have no scientific status.

It often seems that educational discourse draws on the language of Bloom's taxonomy for the discussion of performance and the cognitive capacities that underpin it. It might be more appropriately argued that general educational discourse draws on the same language of logical categories in common parlance that were used as the basis of Bloom's taxonomy. In either case, the language of logical categories (analysis, synthesis, evaluation etc.) seems to have little grounding in student performance in different domains and no relation to the abilities that are explored in psychometrics and cognitive psychology.

The dominance of Bloom-like terminology in educational discourse can be seen in a representative fashion in the notions called the common curriculum elements (CCE) identified by the Queensland Board of Senior Secondary Schools Studies as the underpinning of the Queensland curriculum and the Queensland Core Skills Test (Table 6).

The CCE were identified as a means of grounding in the Queensland curriculum a generic skills test used for between-school scaling of assessment at grade 12. The CCE were developed through an examination of curriculum documents and from extensive consultation with teachers in all areas of the Years 11 and 12 in Queensland. Terms were included in the list if more than two subject areas identified a term as indicating that an activity took place in their subject area. This process is an educational equivalent of kinds of workplace consultation that was involved in the process of defining many work-related skills constructs.

The teachers of a broad range of curriculum areas in Queensland were able to see the terms as indicating some kind of activity that took place in their curriculum area, but there is no evidence that the elements identified have any common or identifiable meaning in different subject areas, or the performance of candidates on the Queensland Core Skills test. It seems that the common curriculum elements are a gesture at what is common to or underpins the curriculum of Queensland, or anywhere else. The process used to develop these common curriculum elements was described by the Queensland Board studies as 'hermeneutic' to contrast it with an empirical or psychometric approach.

As with the Common Curriculum Elements of the Queensland Core Skills Test, curriculum specific discussions of cognitive abilities offers little meaningful definition of what kinds of skills or abilities underpin performance in different subject areas or that are common to much of the curriculum. The CCE are also representative in that they show the limitations of consultation process for determining a theory of ability. It is not adequate to attempt to determine the underpinnings of curriculum or notions of work-related skills from a poll based on unanalysed popular concepts. Thus in Carroll's terms, the Queensland CCE and the Key Competencies have no scientific status.

A theoretically sound (let alone a scientific) theory of abilities cannot be established by referendum on the basis of vague terms. A theory of ability will have to be rigorously conceptualised in carefully defined language, and it may only claim to be a scientific theory if it is explored in terms of actual performances and in terms of the capacities of the mind/brain.

In analysing the notion of ability we have seen that

- the class of tasks approach is limited and incomplete;
- a domain based approach is not generic;
- the psychobiological approach is not yet productive; and
- psychometrics is a kind of constructivist approach that is a provisional but useful synthesis of task, domain and process.

We have also seen that the:

- scientific and empirical study of ability is important;
- psychometric approach to ability claims to be scientific and to have empirical validity; and
- empirical claims of psychometrics are subject to the charge of reification.

The discussion above has been a kind of analysis of and commentary on the zeitgeist of assessment theory and practice at the end of the C20 and the beginning of the C21. It would be useful to move into some particular discussion of the various issues about performance and abilities raised here.

Table 6 The Common Curriculum Elements of the Queensland Senior School Certificate

1	Recognising letters, words and other symbols
2	Finding material in an indexed collection
3	Recalling/remembering
4	Interpreting the meaning of words or other symbols
5	Interpreting the meaning of pictures/illustrations
6	Interpreting the meaning of tables or diagrams or maps or graphs
7	Translating from one form to another
9	Using correct spelling, punctuation, grammar
10	Using vocabulary appropriate to a context
11	Summarising/condensing written text
12	Compiling lists/statistics
13	Recording/noting data
14	Compiling results in a tabular form
15	Graphing
16	Calculating with or without calculators
17	Estimating numerical magnitude
18	Approximating a numerical value
19	Substituting in formulae
20	Setting out/presenting/arranging/displaying
21	Structuring/organising extended written text
22	Structuring/organising a mathematical argument
26	Explaining to others
27	Expounding a viewpoint
28	Empathising
29	Comparing, contrasting
30	Classifying
31	Interrelating ideas/themes/issues
32	Reaching a conclusion which is necessarily true provided a given set of assumptions is true
33	Reaching a conclusion which is consistent with a given set of assumptions
34	Inserting an intermediate between members of a series
34	Inserting an intermediate between members of a series
35	Extrapolating
36	Applying strategies to trial and test ideas and procedures
37	Applying a progression of steps to achieve the required answer
38	Generalising from information
41	Hypothesising
42	Criticising
43	Analysing
44	Synthesising
45	Judging/evaluating
46	Creating/composing/devising
48	Justifying
49	Perceiving patterns
50	Graphicalising
51	Identifying shapes in two and three dimensions
52	Searching and locating items/information
53	Observing systematically
55	Gesturing
57	Manipulating/operating/using equipment
60	Sketching/drawing

5
DeSeCo and the Concept of Competence

It was argued in the preceding chapters that the Key Competencies were developed in a theoretical vacuum, and one can readily agree with critics such as Collins and Stanley (Collins 1993) that the Mayer Committee showed little awareness of and made little use of research into performance and abilities.

The OECD DeSeCo project (outlined in chapter 2), in contrast with the Mayer Committee (and in apparent reaction against the atheortical and inductive empiricism of other psychometrically oriented OECD projects), focussed on the underpinning matters of theory and definition. DeSeCo aimed to establish a clear foundation for defining and selecting key competencies, and a clear foundation for comparing the results of key competencies assessments. To achieve this aim, DeSeCo commissioned experts from various academic disciplines to reflect on the notion of key competencies and to offer from their perspective a set of such competencies. An initial paper was developed for DeSeCo on the concept of competence by Franz Weinert of the Max Planck Institute for Psychological Research (Weinert 2001). Particular papers were also written for DeSeCo by pairs of philosophers and economists, and by a psychologist, a sociologist and an anthropologist (Rychen and Salganik 2001). The outcomes of this work are presented in Table 4 of chapter 2 and the attempt to generalise these outcomes by Rychen and Salganik is presented in Table 5.

Theories of human performance and abilities are not central issues for philosophy, economics, sociology or anthropology. The analysis of theorists writing from within these domains for DeSeCo is not particularly powerful or salient to this discussion or the definition of key competencies. Psychology, on the other hand, is centrally concerned with a theories of human performance and abilities, but the psychologist Helen Haste writes for DeSeCo from the perspective of social rather than cognitive psychology, and her commentary is analytic rather than empirical. The analysis of Haste is not particularly powerful or salient to the definition of key competencies. Franz Weinert, on the other hand, was a cognitive psychologist and his arguments drew on empirical research. His analysis is directly related to the matters at issue here, and in the definition of competencies.

While one must applaud the aim of DeSeCo to deal with issues of theory and methodology, the results of that work shown in Tables 4 and 5 are disappointing. The proposals offered by the various experts are vaguely global or abstract. The most useful set is that offered by the

economists because it is the least ambitious. On the other hand, the economists uncritically accept rather controversial notions like emotional intelligence. It is understandable that the economists would deal with notions like emotional intelligence in an unsophisticated fashion when they are speculating about matters that are not addressed by their specialism. Weinert, on the other hand, is very well acquainted with the arguments in cognitive psychology about abilities, and his responses offers a real test for the notion of key competencies and the arguments at issue here. In reading his paper, however, it soon becomes clear that Weinert is a supporter of certain positions about cognitive performance and abilities.

A Consensually Agreed Core Definition

Weinert was commissioned by DeSeCo to review different theoretical and pragmatic conceptions of competence. According to Weinert, there is no agreed definition of competence:

> Because the scientific definitions of the concept of competence are so heterogeneous, it is impossible to identify a consensually agreed core definition. It is, however, possible to explicate the scientific and practical goals that should be followed with a theoretical construct of competence (p.3).

Weinert's conclusion in the body of his argument is significantly more negative than in the executive summary of the paper:

> The many implicit (in word use) and explicit (in theoretical frames of reference) definitions of competence are so heterogeneous that only a small, vague conceptual core remains (p.26).

The term 'vague' does not occur in the executive summary of Weinert's review. This is an example of the tendency in the whole paper for the summary conclusions to be significantly less sceptical and negative than the analysis and the argument offered in the body of the report. Weinert's analysis shows significant scepticism about the notion of key competencies, but he does not explicitly argue that case. One wonders why he does not draw the conclusion that is evidently suggested by most of his analysis. The conclusion that Weinert does draw (and is drawn by DeSeCo as a whole) is that the definition and selection of key competencies is primarily a political rather than a scientific matter.

Weinert's Definition

Weinert offers the following definition of competence:

> Given the intended use of this report (to provide a conceptual basis for school-based achievement comparisons in international and national systems of reference), it is recommended that competence be considered as a learned, cognitive demand-specific performance disposition, and that corresponding

> metacompetencies and motivational attributions be included in analyses of this construct (p.3).

Weinert's definition is both cognitive and motivational, and one notes the use of the term 'achievement' rather than the generally neutral term ability or the controversial and loaded term 'aptitude' in his definition. Weinert would not use the word aptitude because it focuses on the general abilities that are used to learn new things. Competence in Weinert's definition is learned, presumably to distinguish it from what might be claimed to be innate capacity independent of knowledge. Competence for Weinert is achievement and implicitly, but definitely, distinguished from aptitude.

A crucial element of Weinert's definition is that competence is 'cognitive demand-specific'. Weinert does not explain or elaborate what he means by this term, but it seems to mean the demands that are specific to a particular task. This is, as we will see, consistent with Weinert's emphasis on the place of knowledge in competent or expert performance. Given his emphasis on the demand-specific nature of competence, Weinert has a problem with what is meant by, or can be envisaged as, key competencies because a competency cannot be at once demand-specific and generic.

General Cognitive Competencies

Weinert asserts that psychometric, information processing and Piagetian models of intelligence are general theories of competence:

> The prototypical approaches that focus on general competencies include psychometric models of human intelligence, information processing models, and the Piagetian model of cognitive structural development. Psychometric approaches understand intelligence (competence) as a system of more or less content- and context-free abilities and aptitudes (Carroll, 1993). They provide the cognitive prerequisites for purposeful action, reasoning, successful learning and effective interaction with the environment. In information-processing approaches, intelligence (or general competence) is understood as an 'information processing machine' whose general system features (i.e. processing speed, working memory capacity, processing capacity) allow it to acquire an endless variety of specific knowledge and skills (p.6).

Weinert characterises these general approaches to 'intelligence (competence)' as 'a system of more or less content- and context-free abilities and aptitudes'.

Carroll, cited in the passage above as the representative of the psychometric approach, might justly respond to Weinert's synopsis by stating that the psychometric theories of ability are not necessarily theories of general intelligence, and they are not theories of competence or competencies. Psychometric theories in general, and Carroll's in particular, are theories of differential abilities (Carroll 1993). Unlike the information processing and the Piagetian

approach, the psychometric approach is not necessarily concerned with a notion of general ability. The generally accepted view in the psychometric approach is of abilities as a hierarchy of special, broad and an overarching or underpinning general ability. Carroll is one of the factor analysts who are very circumspect about the importance of a single, general ability. Thus in his comments on general cognitive competencies, Weinert does not give the psychometric approach an adequate review. The next chapter offers such a review of cognitive psychology and psychometric views of abilities.

The important issue in Weinert's analysis of competence is what is meant by 'specialised cognitive competencies':

> Specialized cognitive competencies refer to clusters of cognitive prerequisites that must be available for an individual to perform well in a particular content area (e.g., chess playing, piano playing, automobile driving, mathematical problem solving, trouble-shooting in complex systems, etc.). The domains of specialized competencies can be very narrowly defined (e.g., chess competence) or very broadly and openly defined (e.g., diagnostic competencies).
>
> We know, at least in the case of excellent chess players, that neither general memory capacity, extreme intelligence, nor unusual problem solving skills are necessary for high performance. Rather, the cognitive competence of chess masters and club players, in contrast to beginners, stems from a system of specialized skills and routines, based on thousands of chess configurations stored in memory ('memory chunks'). This is an example of a learned competence, that of course may itself depend on greater or lesser abilities in acquiring this expertise (p.7).

For Weinert, the crucial element of 'specialized cognitive competencies' that distinguishes them from the special abilities of the psychometric tradition is a matter of performance within a 'particular content area'. Weinert offers the familiar examples of chess and diagnostic competencies as a contrast with the abilities developed within the psychometric tradition. Whether they are narrowly or broadly defined, Weinert's specialised competencies are based on the particular knowledge and skills of a domain. They are content and context-based.

Weinert then goes on to refer to the arguments about the nature of expertise in chess, and the way it is based on specially learned knowledge and skills. In referring to 'the greater or lesser abilities' using in acquiring expertise Weinert recognises in passing that the notion that particular expertise is within a particular domain is not at odds with a notion of general abilities. Such general abilities may be the aptitudes that make the development of expertise more or less facile and complete. In some sense Weinert is recognising that notions of general abilities and specific expertise are not mutually exclusive. The issue is whether one

focuses on either aptitude for learning, or expertise. Although these significant distinctions are implicit in some of his comments, Weinert does not explore them.

Work on general abilities within the psychometric tradition would not claim to account for achievements or expertise, and cognitive psychology attempts to analyse the part of performance that can be distinguished from particular knowledge and skills. Cognitive psychology tends to focus on general under-pinning abilities rather than the task-specific aspects of performance. Advocates of the importance of domain-specific knowledge commonly see the importance of knowledge in expert performance as refuting notions of general abilities. Weinert does not explicitly develop such arguments, but he concludes that there are strong theoretical and practical advantages for views of performance specific competencies:

> An overview of research on performance-specific concepts of cognitive competence suggests that this approach has strong advantages over ability-centered definitions of competence because of its theoretical base and pragmatic applications. In particular, the performance-specific approach allows scientific analyses of competence to consider the necessary learning prerequisites for development (p.7).

When he addresses the issue of key competencies, Weinert identifies two promptings for such notions:

> The search for key competences is motivated by two things: (a) the well founded assumption that competencies acquired in school and vocational settings are learned and used in context-specific ways (e.g., within a discipline, within a vocation, within a company), and (b) that most activities over the life course take place in a variety of social and vocational contexts. This has led to the search for context-independent, key competencies that are equivalent in their use and effectiveness across different institutions, different tasks, and under varying demand conditions (p.11).

Weinert states that it is generally assumed that 'competencies acquired in school and vocational settings are learned and used in context-specific ways', but it is not clear what this means, and it is not evidently a common assumption.

In a trivial sense all learning is context specific. Wherever something is learned it is learned in a specific context, and wherever learning is used it is used in a specific context. These things are self-evident. Similarly, all performance is necessarily and inescapably based on knowledge. No performance is content and context free. Again this is self-evident. But there are significant differences between tasks that are recall of propositional knowledge or over-learned and automatised skills on one hand and tasks that emphasise processing and reasoning about unfamiliar information on the other. To emphasise reasoning and the processing of information in an assessment is to focus on those aspects of an individual's performance that are less knowledge dependent and hence more generic and transdomainal.

That is, those tasks that are least familiar and least a matter of specific learning are those where general capacities are most important. (The basis and the implications of these distinctions will be sketched in the Chapter 7.)

The implications of Weinert's claim that 'competencies acquired in school and vocational settings are learned and used in context-specific ways' are not self-evident. It is more appropriately claimed that it is evidently a general assumption that in learning specific things in school and work we are developing the capacity to do other things that we might have to learn in the future. It is generally assumed that learning certain specific things helps one learn other things that are quite different. When we learn something, we can develop capacities that help to deal with quite other kinds of tasks. It is this assumption (and lived experience?) that makes people in general (whatever the views of different theorists) assume and believe in notions of transferable, generic skills.

Weinert sees performance and competence as knowledge and context specific, and this approach is actually at odds with notions of generic key competencies.

The Pedagogical Problem of The Key Competencies

As is commonly the case with critics of generic skills, Weinert readily assumes that giving recognition to (and potentially assessing) generic competencies must entail a certain kind of curriculum and pedagogy (Bowden and Masters 1993; Craigie 1995). On this assumption Weinert draws attention to what he sees as three problems with such a pedagogy. According to Weinert:

> Although the facilitation of key competencies that can be applied flexibly across variable contexts is important and necessary, it has been very difficult to implement this theoretical approach pedagogically (p.12).

Citing his own research, Weinert claims that

> Over the last decades, the cognitive sciences have convincingly demonstrated that content specific skills and knowledge play a crucial role in solving difficult tasks. Generally, key competencies cannot adequately compensate for a lack of content specific knowledge (p.12).

Weinert claims it has been demonstrated that general competencies have virtually no practical utility alone, and that it is specific knowledge, embedded in experience, that is required to successfully implement available competences for solving specific practical problems. On the questionable assumption that giving recognition to key competencies means 'planned instructional programs' or 'special training programs', Weinert states that

attempts to teach matters like critical thinking as such have been challenged as ineffective. He concludes that:

> The definition, analysis and instructional facilitation of key competencies present a scientific dilemma at the present time. From a practical side, it is important that they be fostered; but from a theoretical/empirical side it appears that their learnability as a separate set of competencies is questionable, and is limited to specific content domains (p.12).

Weinert assumes that key competencies are to be learned as a 'separate set of competencies' and then presents arguments that challenge both the meaningfulness of hypothesising, and the practicality of teaching, such competencies. He concludes that from a 'practical side' key competencies are important and must be fostered. It is hard to see how one can conclude that key competencies have to be fostered when they are theoretically and empirically questionable.

Weinert also considers the distinction between curricular and transcurricular competencies, and although he does not recognise a 'theory-based distinction' between them, he recommends maintaining the differentiation between competencies defined as curricular and transcurricular as a 'pragmatic classification'. It seems for Weinert that a pragmatic classification is to be distinguished from a theoretical or empirical one.

Weinert's Notion of Theory and Practice

Weinert claims that an '... analysis of the different definitions and theoretical status of concepts of competence might lead to the inference that the quality of a theory is only weakly correlated with its practical use for educational purposes' (p.22).

One can only wonder at the notion that a high quality theory can be of little practical use for educational purposes. A strong theory is necessarily of practical use and weak theories are in their very nature obstructive, dysfunctional and deleterious. (As the saying goes, there is nothing as practical as a good theory.) One wonders whether what Weinert sees as high quality theory is of little practical use because it cannot offer the advocates of key competencies what they want?

Weinert claims that pragmatically useful models of competence are unlikely to be derived directly from available psychological theories, and we can infer from comments scattered throughout his paper that this is because such theories see, or should see, competence as 'demand specific'. In contrast with the approach of most cognitive psychology, Weinert describes a more promising strategy for developing criteria for a pragmatically useful definition of competence. This strategy would involve analysing different concepts of

competence with the goal of identifying essential theoretical constituents and components. As we saw in Chapter 2, such a pragmatic synthesis seems to be attempted for DeSeCo by Rychen and Salganik.

Without any argument or justification, Weinert concludes that 'no pragmatically useful models of competence can be derived from psychological theories'. This is a remarkably unqualified statement. According to Weinert, psychology, the discipline that has set out to understand human performance and abilities, cannot offer pragmatically useful models of competence.

Weinert phrases the central issue for DeSeCo in the following terms:

> How many theoretical constructs are required to describe and explain nontrivial interindividual differences in performance (p.22)?

This is the fundamental question that is addressed by theories of differential ability in cognitive psychology and psychometric research into individual differences. Weinert concludes that there is little consensus in the social and behavioural sciences about the answer to this question about individual differences in performance. One can reasonably retort, however, that such consensus cannot be expected. A consensus within a field, especially a field focussed on such differences, would be significant in itself. As we will see in the next chapter, there is a good deal of agreement within psychometrics about 'interindividual differences in performance'. This work in psychometrics is worthy of significantly more consideration that Weinert gives to it.

Weinert notes that neurosciences, despite considerable progress, have not been able to approach defining the neurobiological or neuropsychological conditions for cognitive performance or performance differences. Weinert comments in very general terms on research in cognitive psychology, and he concludes that the predictions of performance from intelligence tests are disappointing:

> Looking just at mainstream psychological research, the approach for a long time was to take the concept of intelligence as the most important theoretical basis for predicting and explaining cognitive performance differences. Although results from the thousands of studies were generally in line with expectations, they were also disappointing. Mean correlation coefficients of +.5 (varying between +.1 and +.8) between intelligence test scores and many cognitive performances in school and in work settings have been the rule (p.22).

On the other hand Weinert claims that a combination of general abilities and specific knowledge competencies has proven especially productive in differential and educational psychology for predicting future performance and/or explaining observed performance differences:

> Interindividual performance differences are very stable even in elementary school (correlation coefficients between .6 and .8 between the 3rd and 4th grades). Prior knowledge accounted for over 30% of the variance in performance, and general intellectual ability differences only 10% (as expected this was higher, for example, for mathematical problem solving than for arithmetic tasks) (p.24).

As a result of these theoretical considerations and practical reasons Weinert concludes that 'the concept of competence should be limited to specific knowledge and skills and should exclude intellectual abilities for analysis of school achievement' (p.24).

In his closing suggestions and recommendations Weinert concludes that 'the notion of competence has only a small, vague conceptual core' and that 'a decision about competence concepts needs to be made for achieving scientific and pragmatic objectives': (p.26) He offers two options entitled the conditional and the functional approaches to defining key competencies:

(a) Conditional approach: which cognitive competencies are there?

(b) Functional approach: Which cognitive competencies does one need to master tasks and task demands.

Weinert's conditional approach includes 'intellectual abilities, content-specific knowledge, skills, strategies, metacognitions and action routines', and the approach is said to be exhaustive. It requires an integration of ability and knowledge requisites which, according to Weinert, cognitive psychology has not achieved:

> One would be confronted with a problem not yet solved in the 100-year history of scientific psychology: a complementary classification and performance-specific integration of ability and knowledge. There is neither a theoretical nor a practical solution to this problem at this time. (p.26)

Thus Weinert sees evident difficulties in finding a theoretically legitimate and pragmatically useful answer to the question of how to define competencies. Again he cites his own research to explain the difficulties:

> This can be illustrated by an example from mathematics instruction in the 3rd and 4th grades. Although achievement differences in arithmetic tasks can be almost totally explained by differences in prior knowledge, for mathematical problem solving general intellectual abilities are important as well, although the largest influence still comes from prior knowledge. In addition, with increasing improvement in relevant prior knowledge, there is a decrease in the influence of general intelligence (Helmke & Weinert, 1997).

He concludes that:

> In light of this and similar empirical findings, it is not only justified, but expedient to ignore psychometrically defined general intellectual abilities in favor of specialized knowledge and skills (p.27).

Weinert turns from the matter of cognitive pre-requisites to his preferred functional approach. While we are told that a comprehensive psychological theory of human abilities does not exist, we are also told that the functional approach focuses on the demands of problems and tasks and environmental demands. For some undisclosed reason, the functional approach does not entail or require a comprehensive sociological classification of environmental demands:

> What is required is a prototypical, typical and specific characterization of classes of performance demands, performance criteria and indicators of competencies.

Weinert's position is the class of tasks approach described in the proceeding chapter, but he does not recognise that Carroll, from within the factor analytic tradition, is concerned to empirically establish such classes of task.

Weinert claims that there is good deal of scientific and practical information on which to base such functionally defined key competencies. The characterisation of classes of performance demands is to be based on:

- curriculum theory;
- measurement models;
- task profiles for most vocations, (often defined as a set of action competencies);
- task profiles for typical life situations (economic, administrative and political activities, interaction with mass media, leisure behaviour, travel, social conventions, etc.); and
- task profiles for specific action fields (bank consultant, games such as chess, sales skills, etc.).

This is a very diverse range of sources from which to synthesise key competences, and one can note these sources are not drawn on by Weinert or the other DeSeCo experts. When Weinert references the process of drawing such sources into classes of performance, he refers to nothing more contemporary than The Handbook of Formative and Summative Evaluation of Student Learning by Bloom, Hastings & Madaus published in 1971. Weinert concludes that:

> One can identify specific configurations of cognitive competencies that are necessary for good performance for each of these everyday, school and work

> demands. Primarily, this is a system of learned skills, knowledge, strategies and metacognitions (p.27).

It is disappointing to think that the Bloom Handbook of 1971 is pointing the way forward.

It comes as no surprise when Weinert announces which option he prefers:

> A choice of the second option is recommended. The first option, in which competencies are defined as including both content-free intellectual abilities and learned, content-specific knowledge, confounds two different theoretical approaches: the psychometric model of intellectual abilities and ability differences, and the theory of knowledge acquisition and its use in subsequent learning, problem solving and automatized behaviour (p.27).

In Weinert's preferred position the concept of cognitive competence is limited to learned knowledge, skills and corresponding meta-knowledge. This is a knowledge-centred rather than an ability-centred concept of competence. According to Weinert, a notion of general abilities in a theory of performance is not necessary because differences in general abilities are captured in differences in knowledge acquisition (presumably as measured by amount learned).

While advocating a knowledge-centred concept of competence, Weinert notes that such a conception has one severe disadvantage in that the 'traditional knowledge-centered concept' of competence has been criticised for producing inflated numbers of specific and narrow competencies. He also notes that this proliferation of specific competencies led to the search for key competencies, but he does not go on to consider the relationship between a knowledge-centred concept of competence and the notion of key competencies. Weinert claims that the knowledge-centered concept of competence can be applied to the large number of types of performance demands in daily life, school and work activities, and he claims that they will make international comparisons possible and meaningful. There is nothing in Weinert's discussion that distinguishes his 'knowledge-centred concept of competence' from the international studies of mathematics, science and reading that had already been undertaken by OECD. Weinert acknowledges that there are skills that are useful in many learning situations and for solving different problems:

- mother tongue;
- foreign languages;
- reading and writing procedures;
- mathematical skills; and
- nearly all metacompetencies.

But he states that these key competencies are no substitute for content-specific knowledge, especially when solving more difficult problems. He describes the notion of a few key competencies enabling one to quickly acquire content-specific knowledge as a 'purely utopian view' because only those with enough initial knowledge are able to use new knowledge effectively:

> Thus it is necessary that everyone acquire many important content-specific competencies, including especially narrowly defined key competencies (p.28).

Weinert claims that the 'hypothesised key competencies' cannot be supported by cognitive psychology, and that the key competences have been incorrectly postulated as a breakthrough by many politicians, educators, and journalists:

> In the past, the traditional knowledge-centered concept of competence led to an inflation of more specific, narrow competencies, a frequent source of criticism in curriculum development. A justified response to this was to ask whether there might be a smaller number of key competencies, key qualifications, or basic skills that made it unnecessary to address the acquisition of specific knowledge and that allowed an individual to master new demands. Although this hypothesis could not be supported by cognitive psychology, many politicians, educators, journalists and others still see the postulation of key competencies as a breakthrough. Incorrectly, we must add (see Weinert, 1998) (p.28).

As these key competencies are without theoretical support and empirical meaning, they cannot really be a 'breakthrough', but Weinert describes it as a 'continuous task for the social and behavioural sciences' to 'find compromises that satisfy both theoretical and pragmatic requirements'. It is in this fashion that he envisages the definition and selection of key competencies. The nature of these pragmatic requirements is not analysed or explained by Weinert.

A Response to Weinert's Analysis of Key Competencies

The view implicit in Weinert's discussion is that competence is expertise and expertise is fundamentally a matter of knowledge and experience. Weinert gives little support to the notion of generic skills, and much of his argument is at odds with the notion of key competencies.

Weinert sees the impetus for key competencies coming from managers, bureaucrats and the media, and one can surmise that he actually sees notions of key competencies as naive and perhaps false. Given the necessity to satisfy managers, bureaucrats and the media, Weinert claims that the definition of key competencies should be functional in that it is focussed on classes of tasks rather than a conditional approach focused on what the individual needs to bring to a task. Weinert rejects cognitive psychology as the basis for the definition of

competencies, and he ignores psychometric methodology and the outcomes of psychometric research in particular. Weinert rejects (or more accurately, ignores) the kind of empiricism of the psychometric approach, and he takes a consensual or hermeneutic rather than an empirical stance.

Weinert cannot see how to do anything practical with the notion that competence is demand-specific and knowledge-dependent because competence as knowledge threatens to be endlessly particular. He does not envisage any way in which this multiplicity can be reduced or generalised. As knowledge is task-specific for Weinert, it is not surprising that he comes to the conclusion that competence should be defined in terms of kinds or classes of tasks.

The next chapter will show that arguments that emphasise the importance of particular knowledge in performance are common in cognitive psychology, but it will also show that they are focussed on expertise rather than aptitude, and that they are essentially reactions against notions of general abilities. Notions of generalised abilities developed within cognitive psychology and psychometrics are compatible and consistent with notions of key competencies, but notions of the knowledge specific basis of performance are not. Arguments about the importance of knowledge in competence and performance are refutations of notions of generic skills rather than positive proposals. Because they are refutations, such arguments have to date offered little that is practical or useful.

Whereas Weinert claims that there are two options (an ability/knowledge centred concept and a task centred concept), there are at least three separable issues of ability, knowledge and task. There are distinguishable concepts of competence that view it as either:

- an ability centred concept;
- a knowledge centred concept; and
- a task centred concept.

These separate aspects are not, of course, mutually exclusive options.

Weinert does not adequately deal with the notions of specific, broad and general abilities as developed in the cognitive psychology and the psychometric tradition. Without any justification, he only considers the approach as part of an option where it has somehow to be a theory of knowledge as well as ability. The psychometric tradition has developed theories of abilities, but it is not concerned with the particular knowledge of the expert or with 'competence'.

Weinert eschews cognition as the basis for a set of competencies, and instead he proposes that competencies be defined in terms of classes of tasks. He burdens abilities with knowledge, and then contends that the two are unmanageable and impractical, which only leaves competencies defined in terms of classes of tasks. Weinert does not adequately deal with attempts to understand abilities separate from knowledge, as they are understood in the psychometric tradition. He does not offer a basis for establishing classes of tasks. He has little or no research or theory to point to as a basis for defining kinds of knowledge or kinds of task. And when he comes to identify the basis on which kinds of tasks will be defined and selected he merely cites the work of Bloom in the early 1970s (Bloom, Hastings et al. 1971).

Weinert does not discredit the psychometric approach to the study of abilities, and he is not justified in ignoring it. He does not recognise the fact that psychometrics and cognitive psychology are centrally concerned with general abilities and aptitudes rather than with expertise and competence.

If one agrees with Weinert that expertise and competence are essentially a matter of demand-specific knowledge, one must concomitantly draw the conclusion that notions of generic or key competencies are empty and meaningless, but Weinert does not draw such a conclusion. It is arguable that Weinert should have drawn such a conclusion from his discussion, or that he should have looked more carefully and critically at the notion of competence itself.

A balanced review of the approach of psychometrics and cognitive psychology should have led Weinert to recognise that these research approaches are concerned with general abilities and aptitudes rather than with expertise and competence, and that it is psychometrics and cognitive psychology that offers the best body of research to be consulted in developing and selecting generic or key competencies.

Psychological Approaches to Ability

It has been argued above that the aspiration of psychometrics and cognitive psychology is to identify the kinds of capacities that underpin performance in all sorts of domains. It is in psychometrics and cognitive psychology that one finds an attempt to theorise human abilities and to identify the major aspects or structures of human abilities.

From the beginnings of psychology in phrenology, through the kind of empiricism found in the psychometric approach to cognitive performance, and into the more recent works of cognitive psychology (and cognitive science more generally), there has been an attempt to identify the general kinds of cognitive capacities that underpin performance (Hunt 1993).

The efforts of psychologists to identify and theorise human cognitive abilities is a useful basis (in fact the only basis) for the discussion of the abilities (as distinct from the knowledge and skills of specific domains) that might be and that are assessed in various kinds of activity. We will now turn to this body of theory and research.

6

Theories of Ability from Cognitive Psychology

The previous chapter argued that a concept or construct of work-related skills and abilities, and especially a construct of generic work-related skills and abilities, must entail a tacit theory of ability. A theory of ability in the sense discussed here is an account of, or model of, cognitive performance. Such a model of cognitive performance would (minimally) encompass both what is done and what is needed to do it. A theory of ability or performance becomes scientific in so far as it makes claims or hypotheses that can be supported or refuted with argument and evidence.

The discussion below seeks to outline the development of thinking about human abilities that has taken place in cognitive psychology over the last 100 or so years and to show how some of the key themes in this body of theory shed light on the debate about the Key Competencies outlined in earlier chapters.

The Increasing Interest in Differential Ability Theories

There have been signs of an increasing interest in multiple aptitude and differential ability theories in the educational community in recent years. The popularity of multiple intelligences theory or theories of metacogntion in the 1980s and 1990s expresses a desire to have a conception of cognitive ability and to use such a conception to develop educational practice. Hence the thinking of Howard Gardner (Gardner 1985) about different abilities and that of Robert Sternberg (Sternberg 1985; Sternberg 1997) about cognitive processes has attracted a good deal of non-specialist interest. The more recent interest in notions of emotional intelligence developed by Mayer and Salovey (Mayer 1989; Mayer and Salovey 1997) shows the same desire to expand the notion of intelligence and to theorise ability more broadly.

Such interest in differential ability theories grows from desires to:

- hold a broad and inclusive rather than a narrow view of ability;
- eschew or mitigate assumptions of uniformity and standardisation;
- enrich the notion of what can or should be learned and how it might be learned;
- support more individualised learning and teaching; and

- be more accommodating of individuals and make learning more accessible to all students.

This chapter will examine the work of Gardner and Sternberg in the context of the development of cognitive theory in general. It will also outline some aspects of the debate about IQ between different cognitive theorists.

The Psychometric Approach to Ability

In has been commonly observed by psychometric theorists of ability that the work within the psychometric paradigm has often been casually criticised and dismissed or simply ignored in recent decades. Carroll and Messick, for instance, have complained that the congruence between psychometric theories of ability that developed over the C20 and the theories of Gardner and Sternberg has been ignored, and the strengths of psychometric theorising of ability are not recognised. After dominating thinking about abilities for more than half the C20, psychometric theories of ability seem in more recent times to have been rejected or ignored. This thinking about abilities within cognitive psychology seems to have little impact on popular notions of ability or even the notions of professional educators about abilities. Little of this thinking within cognitive psychology seems to have influenced the thinking about work-related abilities outlined in Chapter 2. The exception is DeSeCo, but this project is in another way a reaction against, and even repudiation of, much of cognitive psychology and psychometrics in particular. An appeal was made to theory and research in criticising the Key Competencies, but those who formulated and those who promoted and defended the Key Competencies did not draw on ability theory or research.

The aim of the discussion in this chapter is to trace the development of theories of ability within psychometrics and cognitive psychology so as to show where the work of Cattell, Carroll, Gardner, Sternberg and Ceci, in particular, fit into this broad body of theory. A second aim of the discussion is to set a context for an assessment of the arguments outlined in Chapter 3 that criticise the notion of generic capacities on the basis of theory and research. A third aim is to show how cognitive psychology might and should inform thinking about cross-curricular and work-related skills and capacities. This last aim involves reviewing the development of psychometrics and cognitive psychology to see what kinds of ability are most usefully posited or theorised as cross-curricular and work-related. Later chapters will show that the kind of theory explored in this chapter shaped the model of performance and the empirical work of assessing the Key Competencies reported in Chapter 9.

Key Themes in Differential Ability Theory

The following questions recur in the theories of ability that are considered below.

- Is there a general cognitive ability?
- What is the nature of and how important is this general ability?
- What more or less distinctly special or primary abilities can be theorised and empirically established?

These questions are clearly cognate with the challenges to the notions of generic capacities outlined in preceding chapters.

- Are abilities transdomainal?
- Is performance, and hence competence, context or domain specific?

In the development of cognitive psychology these issues might be summarised as the debate between the generalists and the specialists.

The Debate Between the Generalists and the Specialists

The major issue in the development of scientific understanding of cognition in the last 100 years has been the attempt to theorise cognitive ability in general and, on the other hand, to discern different cognitive abilities. The issue has played itself out particularly in the interplay between the theories of generalists that give a primary place to a general cognitive ability (commonly called ***g***, the term introduced by Spearman (Spearman 1927)) and the theories of specialists who do not recognise a general ability but rather delineate different special abilities and tend to minimise or deny the significance of overall general ability.

These different views are sometimes summarised as the difference between the British school, initiated by Spearman, that supports a view of a general ability and the America school, initiated by Thurstone, that identified different primary abilities and did not recognise an overall general ability. The same issues are currently being played out in the differences between the more generalist components of Sternberg's triarchic theory of intelligence (Sternberg 1989) and the modular nature of Gardner's multiple intelligences theory (Gardner 1985). Gardner explicitly situates his work within this debate. He describes the issue as a debate between horizontalists who believe in intelligence as a single faculty and verticalists who believe in a set of specialised faculties. He writes of the difference between hedgehogs who seek a single inviolate capacity, that is more or less innate, and foxes who seek an array of human abilities or faculties (Gardner 1985). Ceci's bio-ecological theory of intelligence draws on Sternberg's triarchic theory (Ceci, Ramey et al.

1990), but it differs from that theory in neither recognising a substantial general ability nor transdomainal cognitive processes.

The issues of a unitary or multiple intelligences, metacognition and the notion of generic abilities continues to be important in current educational debates, and what can be learned from the dialectical process of the last 100 years or so in cognitive psychology sheds light on current issues.

The Major Approaches to Understanding Abilities in Cognitive Psychology

Attempting to understand the human mind has been at the centre of the scientific study of psychology that has developed over the last 120 years. A key aspect of this development has been the attempt to understand human learning and intellectual performance, which has in turn prompted thinking about different kinds of ability and the nature of cognitive ability in general. Psychology has been the area of systematic inquiry that explores notions of cognitive performance and notions of generic, cognitive abilities. The beginning of the psychology of cognition has its roots in phrenology, and phrenology was a primitive kind of differential ability theory (Fodor 1983).

Two major streams of development can be identified in the history of psychology over the last century. One stream has attempted to understand the psycho-biological mechanisms of thought and the other stream is more experimental and or theoretical psychology (Sternberg 2000). This experimental and or theoretical stream has developed into cognitive psychology and the more general field of cognitive science in recent decades (Gardner 1987). Another stream has paid particular attention to individual differences in intellectual performance, particularly in the sub-branch of psychology called psychometrics (Carroll 1982; Carroll 1993).

Chapter 4 sketched the way psychometrics was compatible with the behaviourist paradigm that came to dominate psychology in the first half of the C20. The psychometric approach to understanding cognitive abilities was dominant for the first 70 years of the C20. Although the outcomes of their work differed, both Spearman and Thurstone were proponents of this approach.

The 1950s and 60s saw increasingly widespread dissatisfaction with the reluctance of behaviourism to countenance or deal with notions of mental processes. A marked desire to analyse cognitive processes, particularly through the use of computers to model thinking,

developed in this period. These developments in the 1970s came to be known as 'the cognitive revolution' (Gardner 1987).

Psychometrics is the application of techniques of statistical analysis to cognitive performance, particularly the performance of groups of candidates in formal testing situations. The psychometric approach involved the development of techniques for the statistical analysis of test performance. It is sometimes referred to (with unintended irony) as the psychology of individual differences, not because it deals with individuals but because it deals with the differences within groups.

The various approaches to understanding cognition that developed with the 'cognitive revolution' contrast with the psychometric approach either because they are concerned with what is general in cognition (as with Piaget or artificial intelligence research) and are not concerned with 'individual differences', or because they aim to analyse and understand thought processes in general.

The differences between the psychometric and the cognitivist approaches may be summarised as follows.

Table 7 Psychometric and Cognitivist Approaches to Ability

Psychometric	v	Cognitivist
Empiricism	v	Theory
Operationalism	v	Analysis
Experimental	v	Rational
Statistical data	v	Eclectic data
Instrument building	v	Model building

A survey of the last 100 or so years of theoretical and empirical research into human abilities suggests a number of contrasting approaches and lines of development that might be named as follows:

- psychometric or operationalist theories;
- cognitive correlate or 'bottom up' theories;
- information processing or componential theories;
- psycho-biological theories;
- cognitive development theories; and

- socio-cultural or contextualist theories.

The psychometric or operationalist theories focus on the differences between individuals in performance and observable outcomes. In differing degrees this approach is influenced by the behaviourist reluctance to contemplate mental processes in favour of the 'observable behaviour' of test performance. However, in the second part of the C20 psychometrics has become more concerned with attempting to discern and theorise cognitive processes in test performance.

The psychometric approach is shown in the work by Spearman, Thurstone, Vernon, Guilford, Cattell and Carroll outlined below. In his review of the work in this tradition, Spearitt describes this approach as one in which 'the identification of a cognitive ability is based on measures of correlation or other forms of association among cognitive tasks' (Spearitt 1996).

Another approach that developed out of the psychometric approach as part of the cognitive revolution of the 1970s came to be known as the cognitive correlate theories (Sternberg 1994). This work sought to find significant and meaningful correlations between psychometric indicators and performances on various elementary cognitive tasks. This approach can be seen in the work of Pellegrino and Glaser (1979) in which they sought to specify the intellectual abilities that are differentially related to high and low levels of scholastic aptitude. Jensen and others have sought to test the basic processes like the time taken to react to a stimulus (responses that are not seen as matters of intelligence or knowledge) and to then correlate the results of such experiments with intelligence test scores (Jensen 1998). These 'bottom up' theories seek some form of speed of neural processing or nerve conduction as explanations of differences in cognitive ability. The cognitive correlates approach is well represented by the work of Earl Hunt and his colleagues (Hunt, Lunneborg et al. 1975) who asked: 'What does it mean to be high verbal?' and then sought to determine how basic verbal processes like lexical retrieval relate to overall verbal performance on tests of verbal ability. The cognitive correlate approaches are termed 'bottom up' in contrast with the 'top down' approaches of Binet (Wozniak 1999) and Vygotsky (Vygotsky 1978) which are concerned with higher cognitive functions.

The development of theories of information processing in the 1950s and 1960s prompted proponents of the cognitive components approach to ask the question: 'What does the intelligence test test?' This approach aimed at both theoretical and empirical analyses of test performance with the intention of building precise theories of knowledge and cognitive processes. Such work would typically begin with 'task analysis' of test items so as to

identify information processing components, then develop a model of task performance, and then apply such models to research into the differences in the performance of individuals.

Yale psychologist, Robert Sternberg, is the major proponent of the information processing approach, and he began his eclectic cognitive research within this approach. In his Ph.D. research Sternberg sought to decompose analogical reasoning tasks from conventional psychometric tests into separate activities and then relate the separate component activities to overall intelligence (Sternberg 1977). This approach informed Sternberg's triarchic theory of intelligence (Sternberg 1985).

Psycho-biological theories of cognition focus on the relationship between the brain and the mind. They try to bridge the gap between psychology and neurology by attempting to link basic activities of the brain with complex frontal lobe activities. Howard Gardner's multiple intelligences theory is in part a version of this approach (Gardner 1985). The distant but grand goal of cognitive research is to link the 'bottom up' and the 'top down' theories, or the psycho-biological and the cognitive theories.

The cognitive development perspective focuses on the general or typical path of cognitive development of people from birth to maturity. This approach seeks to trace the developmental trajectory of the growing child in an attempt to understand human capacities in general. It seeks the general theoretical principles that govern intellectual growth. Jean Piaget is the initiator and dominant representative of this approach. Howard Gardner was significantly influenced by this approach and he seeks to accommodate aspects of Piaget's work into his multiple intelligences theory.

Socio-cultural or contextualist theories of cognitive performance focus on the relationship between cognition and cultural context. The Russian psychologist, Lev Vygotsky, came to be seen as a major influence on these theories. In contrast with Piaget, Vygotsky saw the development of cognitive abilities as a matter of a person internalising cultural practices rather than a process of the individual accommodating to the world. For Vygotsky learning and intellectual development are largely shaped by a child's interactions with others. These theories look past individuals to consider the way a social context shapes intellectual performance, and they give particular attention to the way culture shapes the development of cognition. Such approaches emphasise the way intelligent behaviour differs in different cultural contexts. These approaches are relativistic rather than hereditarian or nativist. They see intelligent behaviour as socially situated and socially mediated. The bio-ecological theory of Stephan Ceci (Ceci 1990) is a major representative of this approach in cognitive psychology.

The psychometric approach to understanding cognitive abilities was predominant in the first 70 years of the C20. This approach then evolved into the cognitive correlate and component approaches. Developmental theories were prominent in Europe, but they had little impact on cognitive theory in the English-speaking world until the 1970s. The influence of Piaget and the discovery (re-discovery?) of the work of Vygotsky came to have an impact on cognitive theory in the 1970s as behaviourism's influence waned in the English-speaking world. These developments challenged the position of psychometric theories of cognition in the 1970s and in the last twenty five years of the C20 developmental, socio-cultural and psycho-biological theories of cognition came to challenge and rival psychometric theory and psychometric practice. Although psychometric theories of cognition were predominant amongst ability theorists, they have not had much impact on the way educators in general think about learning and cognition or the way the generic skills movement understands work-related skills.

Criticisms of the Psychometric Approach

Criticism of the psychometric approach has become more common and more vociferous in the last twenty five years. The psychometric tradition has been accused of being atheoretical or theoretically weak, and developmental, socio-cultural and psycho-biological theories have gained increasing attention in recent times. Thus there is an increasing acceptance of so-called 'performance assessment' in the United States where psychometric testing is most powerful and the tests of the psychometric tradition are still deeply entrenched in the education system. These tests provoke considerable opposition from some educators and are matters of controversy in the popular media.

The psychometric approach to understanding abilities has been criticised for being:

- based on narrow pencil and paper tasks;
- based on flawed and inconclusive statistical methods;
- merely operational and atheoretical;
- without substantive, real-world meaning and predictive power; and
- culturally biased and, hence, inequitable.

While acceptance of the psychometric approach to cognitive abilities has dwindled in recent years, it remains a substantial body of work, and arguably the most important resource available to inform conceptions of cross-curricular and work-related capacities. Researchers within psychometrics continue to advance the claims of the cognitive theory related to the psychometric approach (Messick 1992, Carroll, 1993 #140). Both Messick and Carroll have not only argued for recognition of the possible contribution of psychometrics to contemporary cognitive theory, but they have also argued that psychometrics can shed light

on and be integrated with developmental, socio-cultural and psycho-biological approaches to cognitive theory.

In some respects it seems that psychometrics is in bad odour. Psychometrics was tainted by its relationship with eugenics and viewed as potentially racist (Gould 1981), and its methods have been roundly criticised with a political passion (Jensen 1980). Even so, the methods of psychometrics are powerful (if limited) and they have significant (if limited) things to say about cognitive performance.

The following discussion in this chapter will outline the development of the psychometric tradition in the C20 and the way it was challenged by the other approaches mentioned above. The strengths and weaknesses of the psychometric approach will be reviewed. Subsequent chapters will argue that the psychometric tradition can shed light on questions about generic skills.

Psychometric Theories of Ability

F. Galton

The history of psychometric theory and practice begins with Frances Galton (1822-1911), the second cousin of Charles Darwin (Hunt 1993; Gardner, Kornharber et al. 1995). Galton was one of the progenitors of scientific psychology, and he had a particular interest in kinds of cognitive ability and differences in such abilities. For Galton, the scientific study of psychology was based on statistical evidence and proof. This commitment to the statistical exploration of test performance is the basis of the psychometric exploration of ability.

Galton can be described as taking a 'bottom up' or cognitive correlates approach to understanding complex cognition. He was particularly interested in different sensory and motor functions and how they related to complex intellectual performances. He studied a variety of psycho-physical tasks, such as weight discrimination and sensitivity to musical pitch, in an attempt to understand complex intellectual performances.

Other researchers followed the example of Galton in exploring such psycho-physical tasks as a means of predicting intellectual ability from basic processes, but by the end of the second decade of the C20 it seemed that little or no relationship between higher cognitive performance and basic processes could be demonstrated.

Galton also worked on complex cognition by comparing the performance of groups of students in different school subjects. To undertake such studies, Galton formulated the correlation coefficient statistic for making comparisons of the performance of students in

different subjects, and hence began the use of statistical correlation (the basic technique of the psychometric approach) for the study of intellectual performance. These techniques were extended and refined by Galton's student and co-worker, Karl Pearson, and the process of statistically analysing cognitive performance that became the psychometric approach was begun.

As well as some striking ideas about cognition, some of the more distasteful aspects of psychometrics can be traced to Galton. It seemed to Galton that genius ran in families and intellectual ability was substantially inherited. He was a convinced hereditarian and believed that a significant amount of an individual's intelligence was genetically determined. He coined the word eugenics and was a major proponent of the eugenics movement that advocated improving the human species by selective breeding (Hearnshaw 1979; Gould 1996).

Galton's studies of the performance of students in different subjects yielded high positive correlations. This fact (or artefact to some) came to be known as the 'positive manifold'. The positive manifold is a significant issue in the theory of cognitive ability.

The Positive Manifold

The work of Galton initiated the long-running, and perhaps eternal debate, between the advocates of a general cognitive ability and those advocating separate, specific abilities. Galton discovered that there was a positive correlation between the school grades of pupils in quite different subject areas. The pupils he studied showed a general pattern of strength or weakness across different subjects, and this prompted him to ask whether there was one or a group of abilities that were involved in different subject areas.

The positive manifold of cognitive performances was found to hold true in almost all instances (on reasonably broad samples of the population) and these observations gave rise to a theory of general abilities or even one general ability that underpins all cognitive performances. On the other hand, it has been argued that the differences between different kinds of cognitive performance are very substantial and that these differences are as important and as informative as the similarities. This second argument has emphasised different special abilities.

The meaning and the significance of the positive manifold and the theory that grows from it has been the subject of on-going dispute. The extent to which a cognitive theory recognises the positive manifold as significant is the key difference between most theories. (See Table 14 slightly adapted from the work of Ceci, comparing different approaches to intelligence.)

The positive manifold was used as the basis of an argument by some for the existence of an overall or general ability or intelligence. Instruments were developed to test this general intelligence and the statistical technique of factor analysis was used by both sides of the debate to support their advocacy of the pre-eminence of either general or special abilities.

A. Binet

The second significant development in the understanding of cognitive abilities took place when the French physician, Albert Binet (1857-1911), was requested to develop a method for testing the intellectual abilities of children after the introduction of compulsory education in France. The purpose of the testing was to determine which children lacked the ability to cope with normal schooling and should be placed in special schools (Hunt 1993; Gardner, Kornharber et al. 1995).

Binet and his co-worker Simon rejected the testing of sensori-motor functions explored by Galton and focussed instead on higher cognitive functions. In contrast to Galton's 'bottom up' approach, Binet aimed to assess complex mental functions from the 'top down'. For Binet judgement was the basis of intelligence. He identified three major elements in intelligent thought: direction; adaptation; and criticism. But Binet's approach was pragmatic in that he sought to discover kinds of task that would differentiate between levels of development without trying to develop a theory as the basis of this differentiation.

Binet and Simon created their first intelligence scale in 1904 to discriminate between normal and mentally deficient children. In 1908 a further version of this scale was developed for differentiating among normal children. Binet's intelligence test assessed skills such as judgment, comprehension, and reasoning, and he sought to use sequences of tasks to define the usual pattern of growth. Unlike the work of Galton, the focus of Binet's work was on more complex mental processes and the tests developed by Binet, unlike the tests of Galton, seemed to successfully identify significant differences between the abilities of individuals. Binet's tests were very successful and they were a major stimulus to the development of standardised testing and the psychometric approach to understanding ability.

Lewis Terman (1877-1956) of Stanford University produced an American version of Binet's tests and he also developed the intelligence quotient (the IQ) as a standardised method of representing relative intellectual ability within and across age groups. The Stanford Binet (as Terman's version of the test was and still is known) began the remarkable development of mass, standardised testing in the United States (Anastasi 1988). The Stanford Binet has undergone some adaptation over the last 80 years, but it remains strongly related to the tests originally developed by Binet.

It seems that the psychometric approach was developed from a synthesis of the statistical methodology of Galton and the test materials of Binet.

C. Spearman

The British psychologist Charles Spearman (1863-1945) was particularly inspired by Galtons' correlational work, and in using the tools developed by Galton he also found that seemingly diverse tests were strongly correlated (Spearman 1927):

> The earliest fundamental observation made was that the inter-test correlation, although widely varying in magnitude, were at least regularly positive in sign (p.7).

Although there was a high correlation between the cognitive tests used by Spearman, there was also a clear pattern or clustering of higher and lower correlations. It seemed to Spearman that there was a kind of hierarchy in the correlations between different tests. In an attempt to reduce the complexity of the correlations between different tests, Spearman made the first steps in the development of the statistical technique called factor analysis that is at the heart of the psychometric tradition (Spearman 1946).

A Note on Factor Analysis

The development of factor analysis began with Spearman's use of the correlation coefficient that Galton had designed for the study of the relationship between the performance of individuals on different tests. Factor analysis developed a complex mathematical rationale and took many forms as it developed throughout much of the twentieth century, but its basic aims and characteristics remain the same. In his Dictionary of Statistics and Methodology (Vogt 1993), Vogt defines factor analysis in the following terms:

> **Factor analysis** Any of several methods of analysis that enable researchers to reduce a large number variables to a smaller number of variables of 'factors' or latent variables. Factor analysis is done by finding patterns among the variations in the values of several variables; a cluster of highly correlated variables is a factor.

One of the industry standard statistical applications for undertaking factor analysis, *The Statistical Package for the Social Sciences*, describes factor analysis in similar terms:

> Factor analysis attempts to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables. Factor analysis is often used in data reduction to identify a small number of factors that explain most of the variance observed in a much larger number of manifest variables (SPSS 9.01).

Gould describes factor analysis as a technique that: simplifies large sets of data by reducing dimensionality and trading some loss of information for the recognition of ordered structure

in fewer dimensions (Gould 1996). Anderson (Anderson 1992) summarises the purpose of factor analysis as follows:

> What factor analysis does is to take a data matrix of intercorrelations and search for underlying traits that could have generated the patterns of co-variance seen in the data matrix (p.23).

Factor analysis is a statistical technique that is designed to achieve the central aim of ability theory's to cluster different kinds of tasks and kinds of thinking into an overall framework.

Spearman and Factor analysis

Through the use of factor analysis, Spearman came to believe there was one significant factor that would explain a large part of the individual differences to be seen in a range of different test performances. Spearman came to view intellectual ability as a kind of 'mental force or energy operating through many specific engines' (Spearman 1927). He defined intelligence as eduction of relations and correlates which involved making inferences about general rules and correlates or the deduction of specific instances from a given general rule.

Spearman developed what is known as the two factor theory which claims that any cognitive performance involves two factors, a universal, general factor that is deployed in any cognitive task and a specific factor for that particular performance. This notion of one large factor of 'general intelligence' (*g*) and a small number of specific factors (*s*) for particular tasks has been very influential. Spearman claimed optimistically that 'the frank adoption of Two Factors would seem to supply an adequate basis – the only one possible – for a unified science of human ability' (p.86).

On one hand, Spearman side-stepped some of the most fundamental and difficult questions about mental abilities in a way that typified much of the psychometric approach:

> We are under no obligation to answer such questions as: Whether 'factors' have any 'real existence'? Do they admit of genuine 'measurement'? Does the notion of 'ability' involve at bottom any kind of cause, or power? or is it only intended for the purposes of bare description? How, if at all, does the role of factors transcend that of classes (p.25)?

Even so, Spearman did not make unguarded claims or insinuations about what had been achieved by the statistical study of test performance:

> In spite of the attempts to bridge the gap between that statistical study of factors and their psychological interpretation, no one can say that a final interpretation of *any* broad factor has been obtained (p.191).

Although Spearman gave emphasis to the general factor, he also identified broad qualitative factors. He introduced the basic verbal and mechanical factor distinction that was developed by Cyril Burt (Burt 1940) and Philip Vernon (Vernon 1973), and he also gave significant support to the notion of social cognition. Unfortunately, Spearman's notion of social cognition was given little attention by most psychometricians until the last quarter of the C20.

L. L. Thurstone

Like Spearman, the American psychologist, L. L. Thurstone (1887-1955) saw great promise in the statistical analysis of test performance using factor analysis (Thurstone 1938):

> One of the oldest psychological problems is to describe and to account for the individual differences in human abilities. How are these abilities and the great variations in human abilities to be comprehended? For centuries philosophers have been free to set up arbitrary classifications of personality types and lists of abilities, and there have been almost as many classifications as there have been writers. The factorial methods have for their object to isolate the primary mental abilities by objective experimental procedures so that it may be a question of fact how many abilities are represented in a set of tasks, and whether a particular objective performance represents an ability that is in some fundamental sense primary (p.1).

Like Spearman, Thurstone looked to the statistical analysis of test performances to shed light on human abilities, but his analysis of test performance led to quite different conclusions. Thurstone developed a method of multiple factor analysis with which he analysed 56 variables into 11 or 12 uncorrelated primary factors. He claimed that 7 or 8 of these factors were clearly interpretable and he saw little evidence in his data of the general ability identified by Spearman. Thurstone identified seven primary mental abilities (Table 8) from his analysis of test performances.

Table 8 Thurstone's Primary Mental Abilities

Word fluency	producing words with a initial given letter, suffix or prefix
Verbal comprehension	knowledge of vocabulary and ability to read
Number	numerical computation and arithmetical reasoning problems
Spatial visualization	mentally manipulating objects in two and three dimensions
Inductive reasoning	identifying rules in word and number series tasks and analogies
Memory	memorising and recalling words and pictures
Perceptual speed	identifying similarities and differences in graphical and numerical material

Thurstone's primary mental abilities was the first comprehensive differential ability construct developed from psychometric analysis, and it is mirrored in more current theories which are outlined in this chapter. As a result of his factor analytic work Thurstone produced the Primary Abilities Test of differential abilities which was widely used between the 1930s and the 1960s. Thurstone's primary abilities influence the Differential Aptitude Tests (Anastasi 1988), the most influential of the multiple abilities batteries currently in use.

Like most psychometric theories of ability, Thurstone's primary abilities are a rather heterogeneous grouping. The first three are kinds of knowledge (related to the notion of crystallised intelligence outlined below). The fourth and fifth are kinds of manipulative processes or kinds of thinking (related to fluid intelligence and hence contrasted with crystallised intelligence). The second last of Thurstone's abilities is a basic cognitive function, and the last is an outcome or characteristic of the functioning of other processes.

P. E. Vernon

A major step in the development of cognitive ability theory was taken by British psychologist and factor analyst Phillip Vernon (1905-1987). Vernon was the first to attempt a full scale reconciliation of the positions of Spearman and Thurstone (the generalists and the specialists factorists) by developing an hierarchical view of the structure of abilities.

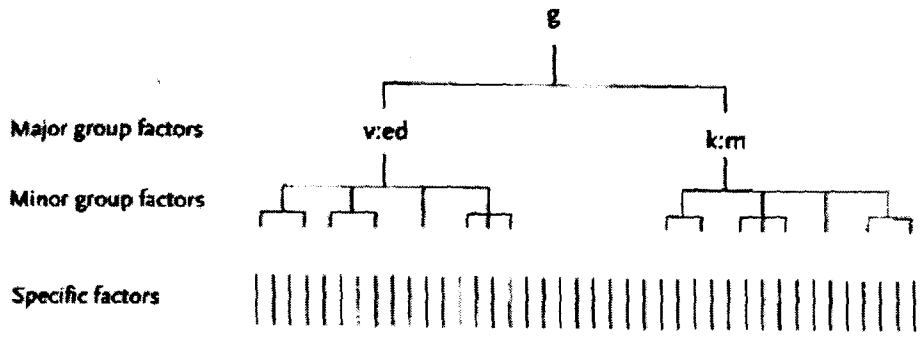


Figure 5 Vernon's Hierarchical Model of Intelligence

In his book The Structure of Human Abilities (Vernon 1973), Vernon reviewed the results of factor analytic studies to that time. He identified a general intellectual factor and two major group factors:

- **v:ed** or verbal/educational ability, and
- **k:m** or practical/mechanical ability.

He also identified various minor group factors and specific factors that he organised into an hierarchical model that envisaged a general ability which correlated with the specific and the group factors.

Although Vernon recognised the importance of factors other than ***g***, he saw ***g*** as accounting for a significant amount of the differences in test performance. He believed 'most of the variance of human abilities in daily life is due to ***g***', and he also believed ***g*** was more important than specific aptitudes in job success.

Vernon took a rather atheoretical and behaviourist approach to defining and testing cognitive abilities. Where Thurstone used psychometric constructs to develop his theory, Vernon based parts of his work on areas of academic study and skill. The generic skills of reading and handwriting were included in Vernon's model, but so were specific subject areas. Vernon's model of abilities placed **V:ed** at the centre of a constellation abilities, and closest to ***g***.

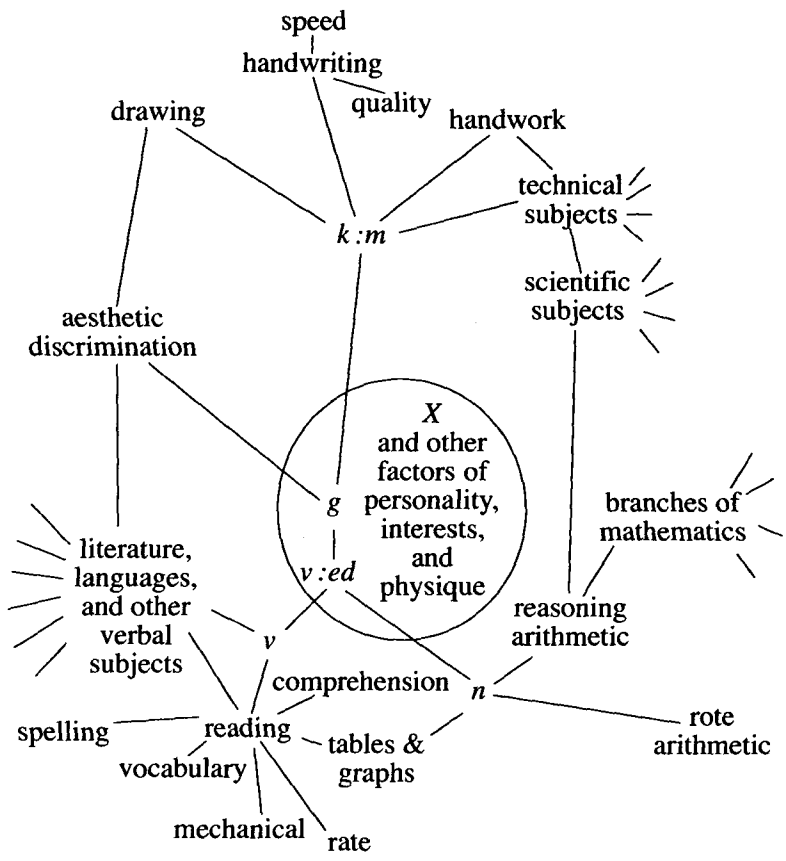


Figure 2 Vernon's Structure of Educational Abilities (p.47)

The work of Cattell moved, on the other hand, away from a notion of ***g*** at the centre of **V:ed**.

R. B. Cattell

A more sophisticated reconciliation of the contrasting views of Spearman and Thurstone was produced by Raymond Cattell (1905-1998) and John Horn in the early 1970s (Cattell 1987). They developed an hierarchical theory (known as **Gf-Gc** theory) which identified two major types of general factors and three minor ones.

Cattell called his position a 'triadic theory' which was made up of capacities, provincial powers and agencies. Cattell's theory is essentially a two stratum one including an hierarchical combination of general and special abilities.

In Cattell's theory capacities are abilities reflecting limits to brain action as a whole; provincial powers are types of local organization for different sensory and motor modalities; and agencies are abilities to perform in different areas of cultural content, acquired through the investment of fluid intelligence in learning.

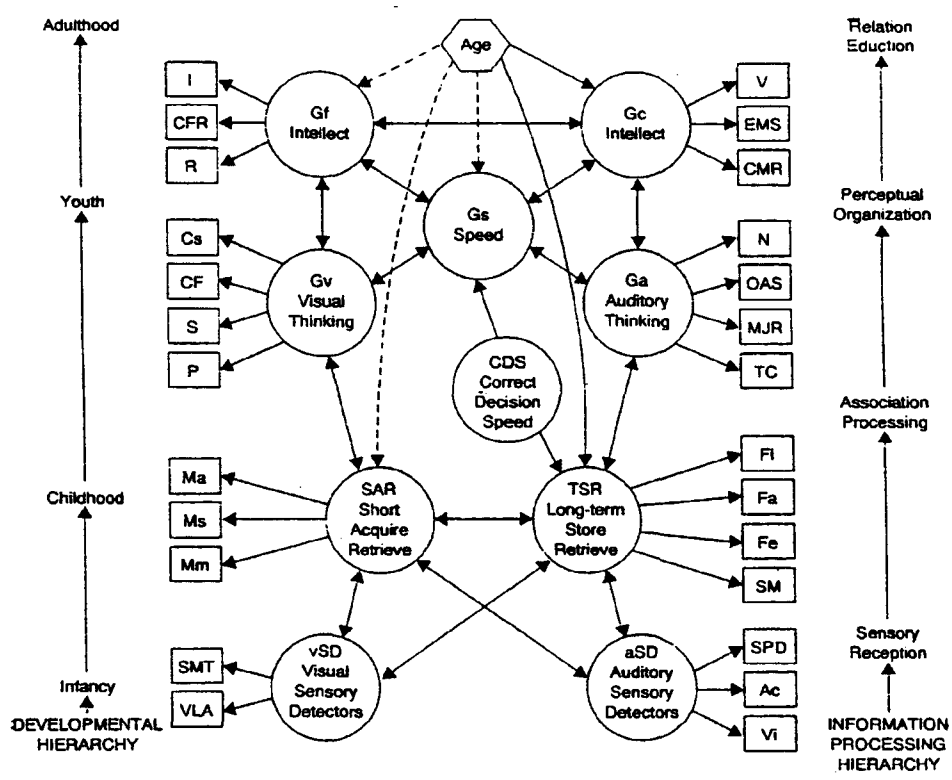


Figure 3 Cattell's Triadic Theory of Intelligence (Encyclopedia of Intelligence p.444)

Cattell also made an influential distinction between two kinds of intelligence which he described as 'crystallised' on one hand and 'fluid' on the other. According to Cattell, fluid intelligence is a flexible capacity that deals with novel challenges and crystallized intelligence is the ability that is developed over time as a result of experience. Cattell saw crystallised ability as a result of the 'investment' of fluid ability.

In general terms, Cattell offered the following breakdown of intelligence.

Table 9 Cattell's Gf-Gc Theory of Intelligence

Gc	verbal knowledge, reading comprehension, and prior educational achievement (similar to Verbal-educational factor of Vernon) required for school activities
Gf	abstract, non-verbal reasoning that is a basic biological capacity that can be measured as perceptual ability (quantitative and other kinds of reasoning relate to this factor)
Gv	spatial relations, visualisation (related to the practical-mechanical ability of Vernon)
Gr	memory retrival
Gs	performance speed

According to Cattell, verbal tests involved crystallized intelligence and what is commonly meant by 'general intelligence' is best seen in verbal intelligence tests such as the Stanford Binet or the Weschler intelligence tests (Anastasi 1988). On the other hand fluid intelligence is elicited by non-verbal tests such as Cattell's own culture fair test (Cattell 1987) and the Progressive Matrices developed by Raven (Anastasi 1988) to test Spearman's ***g***. Fluid abilities are the reasoning and problem-solving capacities measured by such psychometric test items as the analogies, classifications, and series completions. Crystallized abilities can be said to derive from fluid abilities and can be viewed as their products. Tests of crystallised ability include vocabulary, general information, and knowledge about specific fields.

In Cattell's model, the omnibus view of ***g*** typified by the Stanford Binet or the Weschler intelligence scales has given way to an abstract, non-verbal reasoning of the kind found in Raven's Progressive Matrices. (See appendix 3.)

Cattell attempted a comprehensive theory of cognitive performance that included a view of development, a view of basic information processing, a view of higher order abilities and a view of general abilities. Cattell's hierarchical model has been influential, particularly the distinction between crystrallised and fluid intelligence. According to Carroll, Cattell's analysis of abilities 'comes very close to being a cognitive analysis' (Carroll 1993).

J. P. Guilford

Like Spearman, Thurstone, Vernon and Cattell, J. P. Guilford (1897-1988) used factor analysis to explore human abilities, but unlike the earlier factor analysts Guilford based his work on a taxonomic model of intelligence formed by a three dimensional matrix of operations, products and contents. In The Nature of Human Intelligence (1967) Guilford claimed his model offered a link between psychometrically based theory and a general psychological theory

Guilford took what he termed a 'multivariate view of intelligence' (Guilford and Hoepfner 1971) in opposition to Spearman. He found negligible between test correlations for 15 to 25% of cognitive tests. Guilford was particularly aware of differences of ability within persons:

> In an ordinary kind of population, Bloom (1963) found that if we were to define as 'gifted' a child who is in the highest 10% on any one of Thurstone's PMA tests, as many as 60% could regarded as gifted (p.28).

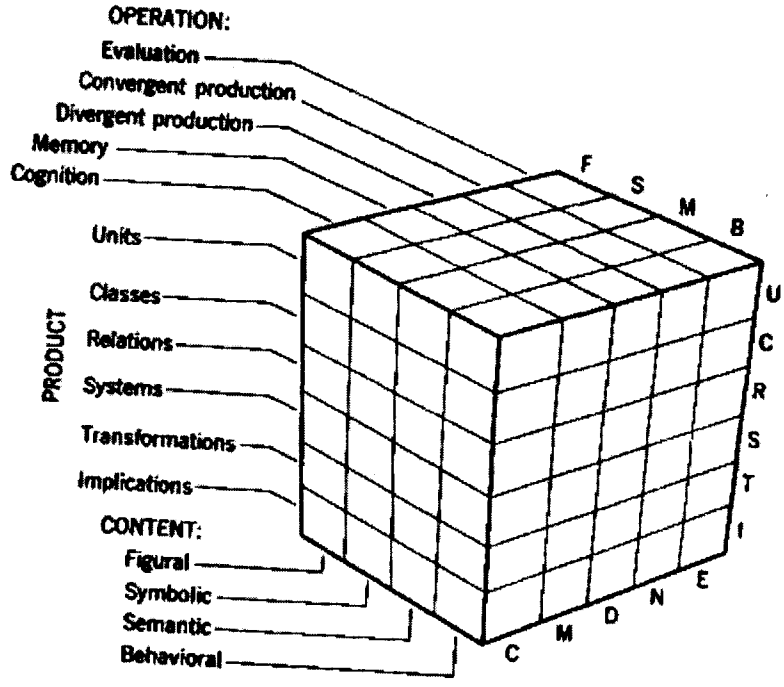


Figure 6 Guildford's Structure of Intellect Model

For Guilford, a 'psychological factor' is demonstrated by a group of tests that are correlated and it is an underlying, latent variable along which individuals differ. Guilford aimed to 'tie factors to a psychological theory' in which the factors would be 'isolable psychological functions' (p.43). For Guilford: 'All learning generalises somewhat: transfer is the rule, not the exception' (p.43).

In his structure of intellect theory published in 1967, Guilford proposed an *apriori* model on three dimensions made up of operations, content and products.

Table 10 Guilford's Structure of Intellect Model

Operation	Content	Product
Cognition of	Figural	Units
Memory for	Symbolic	Classes
Divergent production of	Semantic	Relations
Convergent production of	Behavioural	Systems
Evaluation of		Transformations
		Implications

The structure of intellect theory was an elaborated version of Thurstone's Primary Mental Abilities. Like Thurstone, Guilford found no evidence of a single general intelligence. He hypothesised some 120 distinct abilities and sought to empirically demonstrate the separateness of many of them:

> If one demands a *g* factor, he can always have one. But it would seem illogical to insist upon it in the face of zero correlations (p56).

In his factor analytic research, Guildford aimed to develop empirically based taxonomic concepts and to empirically validate the different aspects of the structure of intellect theory. He claimed to have confirmed 98 of the 120 factors proposed by his model.

To some it seemed that Guilford exploded the factor analytic enterprise by an extreme multiplication of factors, and created what came to be know as 'bloated particulars ' (Kline 1991) or the 'problem of factor proliferation' (Rowe 1985). He created a differential ability theory that is so diffuse that it loses its usefulness. Although it seemed very promising in the 1960s, in the 1990s Carroll dismissed the work of Guilford as 'a somewhat eccentric aberration in the history of intelligence models' (Carroll 1993).

The Cognitive Revolution

At the time of Guildford's factor analytic work, a quite different approach was gathering momentum under the influence of multi-disciplinary cognitive science and artificial intelligence (AI) research (Gardner 1987).

Allen Newell and Herbert Simon developed innovative and influential computer models of cognition, and it seemed that AI was about to shed significant light on intelligent performance (Gardner, Kornharber et al. 1995). A range of researchers sought to use computers to solve problems and to learn. These general problem solving algorithms and heuristics that were built up through AI research had impressive results with chess and mathematical problems. In time, however, the inability of AI machines to reason with informal commonsense or to process natural language made the differences between AI and human thought glaringly evident.

The importance of a broad and rich knowledge base to human cognition was raised to prominence in the late 1970s and 1980s because of the limited success of the generalised algorithms used for problem solving in AI. Although there were impressive developments with neural networks and parallel distributed processing during the 1970s and 1980s, optimism about understanding human cognition through AI receded (Gardner 1987). By the 1980s it seemed that the information processing approach to higher cognitive functions

could not find a model for human thinking in current computers, and the distinctive characteristics of human thought remained elusive (Das, Kirby et al. 1975).

During the 1970s and 1980s criticism of the psychometric approach led to increased efforts to understand cognitive processes as parts of a coherent explanation of cognition. Under the influence of Russian theorist Luria, Das, Kirby and Jarman (Das, Kirby et al. 1975) developed a theory of cognitive processes. They tried to relate mental ability to neurological function by developing a notion of simultaneous and successive modes or styles of cognitive processing. This work influenced the intelligence tests produced by Kaufman in the 1980s (Anastasi 1988).

The cognitive correlates approach taken by researchers like Earl Hunt (Hunt, Lunneborg et al. 1975; Hunt 1978) involved a 'bottom up' analysis and understanding of cognition. Such researchers looked for correlations between intelligence test scores and basic cognitive activities like speed of reaction to stimulus. Taking the same approach, Arthur Jensen used the most basic tasks that involved little knowledge or information and tried to determine the extent to which processing speed and efficiency are related to intelligence (Jensen 1998).

J.B. Carroll

The development of the factor analytic tradition peaked with the publication of John Carroll's Human Cognitive Abilities: A Survey of Factor-Analytic Studies (Carroll 1993). In this compendious work, Carroll undertook a consistent reanalysis of 461 factor analytic data sets, the most important data sets produced between the 1930s to the 1990s. His aim was 'to present what is known about cognitive abilities and its scientific basis' (p.28).

For Carroll, the factor analytic approach is a matter of 'empirical studies of ability – a science of human abilities'. He describes his enterprise in the following terms:

> The investigations dealt with in this book can be regarded as attempts to identify abilities by systematically classifying different tasks with respect to the abilities they appear to require (p.9).

and

> A factor, if it is well established in a number of empirical investigations, is in essence a latent trait reflecting differences over individuals in ability characteristics or potentials (p.20).

The factors defining abilities at each of the three levels of his model are observable differences in the performance of individuals on particular classes of tasks.

Carroll sees a fundamental unity in the psychometric attempts to define ability:

> All models of the structure of cognitive abilities appear to be attempts to classify the various manifestations of mental ability according to aspects of content, type of processing required and type of response or outcome' (p.67).

As a result of his re-analysis, Carroll produced a three stratum hierarchical theory according to which cognitive abilities are either narrow, broad or general:

> There exist a substantial number of distinguishable and important mental abilities, as many as 30 or more. While it may well be the case that general intelligence, a recognised higher-order factor of cognitive abilities, is the most weighty element in all these relationships ... the possible importance of more specialised abilities cannot and should not be ignored (p27).

The three levels of Carroll's theory are:

- the first level of narrow abilities that reflect experience or learning, and are related to Thurstone's primary abilitites;
- the second level of broad abilities which are moderately specialised and differ in terms of process, content and manner of response; and
- the third level of a general cognitive ability.

Carroll accepts Spearman's notion of ***g*** and says his results are consistent with Spearman's two factor model. But Carroll does not see fundamental differences between the work of Spearman and Thurstone, and his model incorporates the primary abilities of Thurstone at the first stratum.

According to Carroll, the work of Spearman and Thurstone can be satisfactorily reconciled in a hierarchical model so that they only differ in the 'relative importance they attributed to primary and second order factors' (p50). He notes that Thurstone originally developed a one stratum theory but that in later years Thurstone became interested in including group factors and a second order general factor:

> Thurstone's assessment of the importance of group factors seems to be justified (p.57).

According to Carroll, many aspects of Vernon's hierarchical model are correct:

> There is good evidence, for example, for clustering of variables around higher-order verbal-educational and spatial-mechanical factors, and for the domination of these factors by some sort of general factor (p.60).

Carroll describes the Cattell-Horn model, as summarized by Horn (1985, 1988), as

> ...a true hierarchical model covering all major domains of intellectual functioning. Numerous details remain to be filled in through further research, but among available models it appears to offer the most well-founded and reasonable approach to an acceptable theory of the structure of cognitive abilities. The major reservation I would make about it is that it appears not to provide for a third-order *g* factor to account for correlations among the broad second-order factors (p.62).

Carroll also claims that there is a fairly close correspondence between his view of intelligence and Gardner's theory of multiple intelligences (p.641). Carroll identifies 20 domains of cognitive ability in which he groups factors:

> That is, the first nine of these groups are regarded as true cognitive abilities in the sense of being relatively fixed, long-term attributes of individuals respecting the kinds of cognitive tasks they can and cannot perform, with varying degrees of success, at a particular stage of development (p.137).

Table 11 gives a general outline of the three-stratum structure of the major cognitive abilities that have been identified in Carroll's survey. According to Carroll, his analysis of abilities of several orders and strata offers insight into the structure of abilities and can be the basis for a scientific theory of cognitive abilities.

In Carroll's model a factor G (3G) at Stratum 3 dominates a series of broad abilities at stratum 2. Carroll finds consistent evidence of a general factor at Stratum 3 of his model:

> There is abundant evidence for a factor of general intelligence, G (or 3G), found at the highest order (usually 2 or 3) of analysis for a given data set and thus at stratum III, that dominates factors or variables that emphasize the level of difficulty that can be mastered in performing induction, reasoning, visualization and language comprehension tasks. There is also some evidence that the G factor is likely to be correlated (though at a low level) with measures of speed of information processing and capacity of working memory (p.622).

Carroll warns that inadequacies of available data mean it is not advisable to attempt to assign numbers or coefficients to indicate how strong this domination by Stratum 3 is, in the case of each second-stratum factor. According to Carroll, the strength of domination is roughly indicated in the table by the closeness of each box representing a second-stratum factor to the left hand margin. This places fluid and then crystallised intelligence as closest to general intelligence. At a lower order of analysis, or at Stratum 2, a number of broad ability factors can be distinguished:

- language:
- memory and learning;
- visual perception;
- information-processing; and
- knowledge of certain general domains.

Table 11 Carroll's Three Stratum Structure of Cognitive Abilities

Stratum 3	General Intelligence 3G							
Stratum 2	**Fluid intelligence 2F**	**Crystallised intelligence 2C**	**General memory and learning 2Y**	**Broad visual perception 2V**	**Broad auditory perception 2U**	**Broad retrieval ability 2R**	**Broad cognitive speediness 2S**	**Processing speed 2T (RT decision speed)**
Stratum 1	**Level factors** General sequential reasoning (RG) Inductive (I) Quantitative reasoning (RQ) Piagetian reasoning (RP) **Speed & level factors** Speed of reasoning (RE)	**Level factors** Language development (LD) Verbal (printed) language comprehension (RC) Lexical knowledge (VL) Reading decoding (RD) Cloze ability (CZ) Spelling ability (SG) Phonetic coding (PC) Grammatical sensitivity (MY) Foreign language aptitude (LA) Communication ability (CM) Listening ability (LS) Foreign language proficiency (KL) Speed & level factors Reading speed (RS) Oral production and fluency (OP) Writing ability (WA)	**Level factors** Memory span (MS) **Speed & level factors** Associative memory (MA) Free recall memory (M6) Meaningful memory (MM) Visual memory (MV) Learning ability (L1)	**Level factors** Visualisation (VZ) **Speed & level factors** Spatial relations (SR) Closure speed (CS) Flexibility of closure (CF) Serial perceptual integration (PI) Spatial scanning (SS) Perceptual speed (P) **Miscellaneous** Imagery (M) Length estimation (LE) Perception of illusions (IL) Perceptual alternations (PN)	**Level factors** Hearing & speech threshold (UA, UT, UU) Speech sound discrimination (US) General sound discrimination (U3) Sound frequency discrimination (U5) Sound intensity duration discrimination (U6) Musical discrimination & judgement (U1, U9) Resistance to auditory stimulus distortion (UR) Temporal tracking (UK) Maintaining & judging rhythm (U8) Memory for sound patterns (UM) Absolute pitch (UP) Sound localisation (UL)	**Level factors** Originality/ creativity (FO) **Speed & level factors** Ideational fluency (F) Naming fluency (NA) Associational fluency (FA) Expressional fluency (FE) Word fluency (FW) Sensitivity to problems (SP) Figural fluency (FF) Figural flexibility (FX)	**Speed factors** Rate of test taking (R9) Number facility (N)	**Speed factors** Reaction time (R1) Choice reaction time (R2) Semantic processing speed (R4) Mental comparison speed (R7)

The stratum I factors listed under each second-stratum factor of Carroll's model are those regarded as being most likely to be dominated by the respective stratum 2 factor.

Although Carroll develops a hierarchical model of cognitive abilities that includes a concept of ***g***, he is agnostic about its impact or the significance of this general factor. Other theorists have continued to advocate the importance of a concept of ***g*** or IQ, but the notion of ***g*** and IQ has also drawn significant criticism.

In reviewing Carroll's work Spearitt (Spearitt 1996) sees significant implications in the model for teaching and learning including the:

- identification of learning difficulties;
- adaptive teaching; and
- training of stratum 1 abilities.

Spearitt also sees in Carroll's model implications for research and curriculum development and the organisation of classrooms, schools and school systems.

A Response to Carroll

The work of Carroll is significant and deserving of attention, but it remains rather opaquely technical, and it seems to have had had little or no impact or influence on educators, let alone the public in general. Rather too much of Carroll's analysis is of theoretical rather than practical interest, and it isn't given real world meaning and relevance. While they may account for differences in performance, stratum 2 factors of broad memory, perception, retrieval and speed do not have much practical meaning, or if they do, Carroll does not communicate the significance of these factors.

Carroll's tireless analysis seems to add up to the few significant findings that there are three key abilities: verbal, abstract reasoning and spatial/visual abilities. Carroll also endorses Cattell's notion of fluid and crystallised abilities, where abstract and spatial/visual abilities are the most fluid, and verbal ability is the most crystallised.

Factor Analytic Views of Abilities

Factor analytic research has distinguished three key factors of verbal, abstract/logical and spatial/visual. All factor analytic theories have a concept of verbal ability. For Carroll verbal ability is the residual after removal the fluid reasoning ability. A non-verbal, abstract and logical reasoning ability is consistent across the psychometric models. A non-verbal spatial visualisation ability is also consistent across the models. The distinction between fluid and crystallised intelligence has been very influential.

The weakness of factor analytic research is in the extent to which it is based on laboratory tasks and pencil and paper tests. Similarly, the constructs identified by factor analysis are not readily related to real-world activity.

While it might seem that factor analytic research adds up to a fairly modest return for a great deal of effort, it is none the less substantially grounded and deserving of attention. However when one compares the results of factor analytic research with the notions of work-related skills outlined in Chapter 2, one sees that psychometric notions had little or no impact on these conceptions. The way that psychometric conceptions and common parlance conceptions might be related will be examined in subsequent chapters.

The psychometric approach to defining cognitive abilities has been subject to considerable challenge and debate. The arguments advanced by supporters of ***g*** or IQ, like Jensen (Jensen 1998) and Gottfredson (Gottfredson 1997; Gottfredson 1998), have excited scepticism from Sternberg and active repudiation by Gardner and Ceci. Arguments about the nature of cognitive abilities readily extend into arguments about the influence of cognitive abilities on various positive and negative social outcomes. Defining and assessing cognitive abilities has historically been involved with political arguments about differences in race, gender and class (Herrnstein and Murray 1994).

Before moving on to the views of these theorists who challenge factor analytic approaches to cognitive abilities and the notion of IQ, it would be appropriate to mention the developmental approach to understanding cognition of Piaget.

The Developmental Perspective

After some initial stimulus from Binet, the biologically trained Jean Piaget sought to trace the pattern of human intellectual development from the initial and rudimentary to the sophisticated. Piaget focussed on the development of logical reasoning, and was concerned with general principles rather than individual differences (Piaget 1971).

As a biologist, Piaget was particularly concerned with interactions between the organism and the world, but his concern with the structure of thinking and the mind contrasts with socio-cultural schools of thought. As a structuralist, Piaget was most interested in general or universal patterns and, in contrast with Vygotsky, he had comparatively little to say about, or gave comparatively little emphasis to, conceptualising or exploring theories of differential ability or the impact of different social contexts on the development of cognition.

Piaget was concerned with the general trajectory of cognitive development. He saw development taking place in giant waves as a results of the two fundamental processes of assimilation and accommodation. Piaget identified four stages of development:

- sensori-motor period;
- pre-operational;
- concrete operational; and
- formal operational.

Development went from sensori-motor actions to concrete, and then on to formal operations.

Although the work of Piaget has been subject to significant and telling criticism, it has continued to find supporters and adapters (Gardner 1974). Gardner was much impressed by the work of Piaget, and he saw it as a significant account of the development of logico-mathematical intelligence. Ceci recognises that Piaget is fundamentally at odds with the opposition to a unitary intelligence that he and Gardner share (Ceci 1990):

> Piaget's theory has long been praised and criticized for the same reason: It depicts intellectual growth through a series of stages and resultant structures that are genetically programmed to unfold in response to what are normally thought to be universal experiences, in invariant order, although not invariant rate. Piaget saw development as a 'structured whole,' with a singular, integrated intellect always orchestrating all of one's cognizing, regardless of the specific domain involved (moral, spatial, linguistic, aesthetic, mathematical, or whatever). Piaget's theory assumes that the structures that comprise the intellect are powerful 'transdomainal' algorithms that are 'blind' to content: Once a logico-mathematical structure is in place, it can be recruited by any task that requires it (p.198).

What Ceci identifies as the significant characteristics of Piagetian theory (that of being transdomainal and blind to content) were the key grounds for the theoretical argument against the Key Competencies reviewed in Chapter 3. Challenging such claims is a key element of the work of Ceci, and, in some respects, it is also a key element of the work of Gardner.

There have also been other reactions against the psychometric approach to understanding cognitive abilities.

Information Processing or Componential Theories

Robert J. Sternberg of Yale University is the pre-eminent representative of the information processing and componential approach to understanding intellectual abilities. He would style himself as a cognitive psychologist concerned to understand intellectual processing. His work grows out of, but involves consistent criticism of, the psychometric tradition (Sternberg 1985; Sternberg 1996). Sternberg has worked in psychometric test construction and he usually seems to share the assumptions about research of the psychometric tradition. He has attempted to create psychometrically based commercial tests that grow from his theorising about intelligence (Sternberg 1991).

The information processing approach of Sternberg is concerned with general cognitive processes and hence is consistent with a notion of general intelligence. His early work focuses on the way general processes operate at different levels and it gives little attention to knowledge, or the content of thought, or to the context of thought. Over time however, Sternberg has become an increasingly vocal and trenchant critic of the refusal or inability of the psychometric approach to explain intellectual processes. During the 1980s Sternberg aimed to develop a comprehensive, explanatory theory of intelligence.

Sternberg's initial PhD research (Sternberg 1977) aimed at analysing the cognitive processing required by conventional test items, particularly analogical reasoning items, to give a process analysis of test performance. This work involved the decomposition of analogy tasks into components and then an attempt to see whether and how different component processes related to overall performance. Sternberg analysed analogy tasks into the general processes of:

- encoding;
- inference;
- application;
- justification; and
- preparation-response.

It is this decomposition of test tasks into component processes that can to be termed the cognitive components approach.

Sternberg was alert to growing criticism of the notion of general, or unitary, intelligence and other aspects of the psychometric tradition. His approach was eclectic as he aimed to synthesise a unified or unifying theory drawing on a wide range of evidence to explain a wide range of mental activity. Sternberg also became increasingly sceptical about the notion of innate intellectual capacity, and manifested this scepticism in various attempts to demonstrate the plasticity of intelligence by improving intelligence test scores through instruction (Sternberg 1983).

Sternberg felt the need for a theory of intelligence that could adequately describe cognitive processes and give recognition to the importance of experience and culture in intellectual performance. He created such an account in his triarchic theory (Sternberg 1989).

Sternberg's Triarchic Theory

Sternberg's triarchic theory of intelligence is a broad and comprehensive, but not a complete, view of cognitive performance. As indicated by the name, the theory has 3 sub-theories: the contextual; the experiential; and the componential.

The contextual sub-theory is concerned with the way intelligent behaviour is defined by the sociocultural context in which is situated. This involves adaptation to the environment, selection of better environments, and shaping of the present environment. This anthropological aspect of the theory is mildly relativistic, but the emphasis Sternberg places on the directed nature of intelligent thought contrasts with the views of strong contextualists mentioned below (Collier 1994).

The experiental sub-theory deals with the relationship between intelligent behaviour and the experience of the individual. It proposes that intelligent behaviour be interpreted along a continuum of experience from novel to highly familiar tasks/situations. It also proposes that performance changes as repeated interactions move tasks from being relatively novel to being familiar and then automatised. Sternberg proposes that intelligence is best measured with relatively novel tasks that involve the need to automatise information processing.

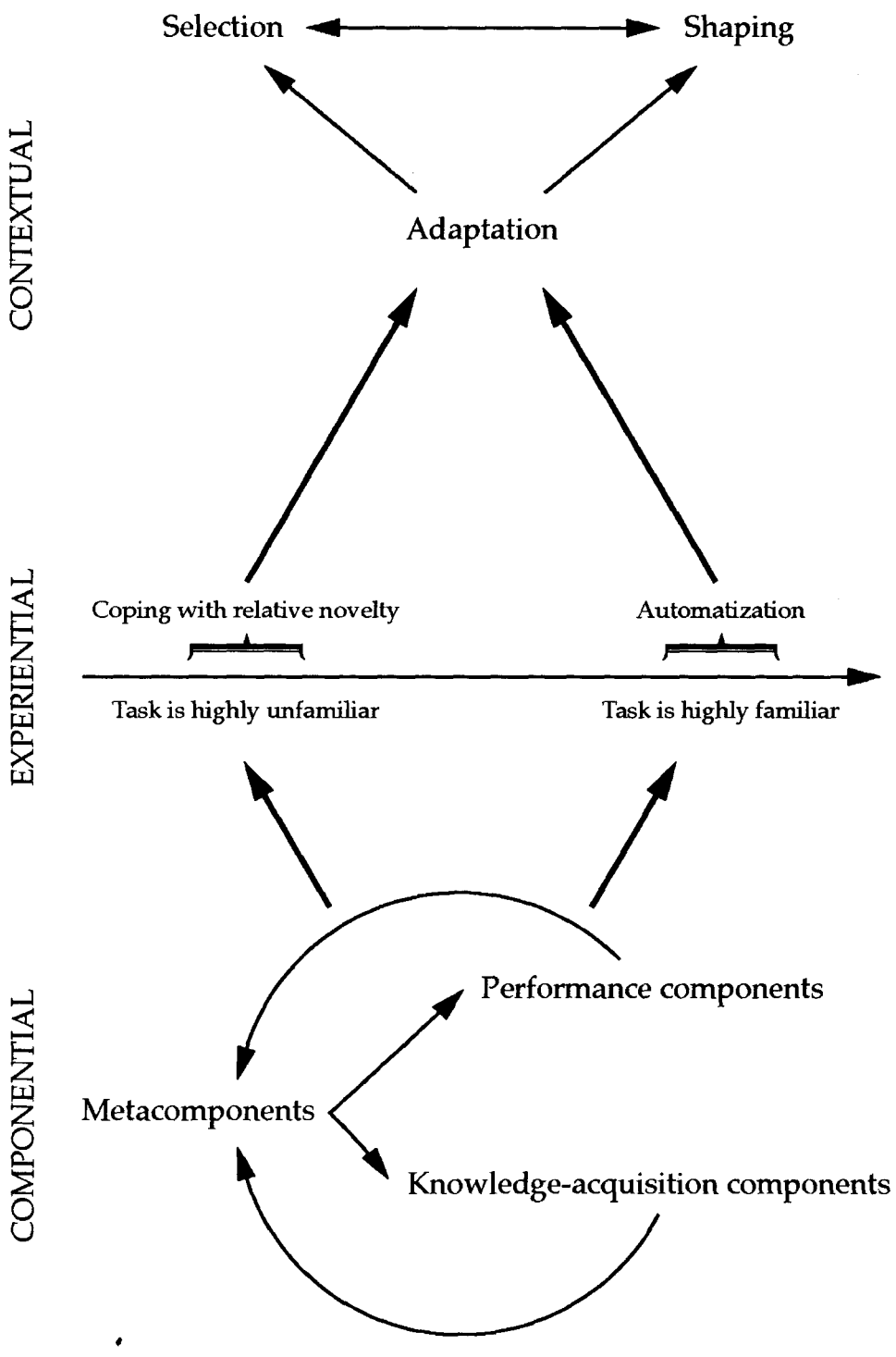


Figure 7 Sternberg's Triarchic Theory of Intelligence (Sternberg 1989)

Under this sub-theory Sternberg identified 'three main factors or constellations of skill':

- problem solving ability;
- verbal ability; and
- social competence.

Sternberg has become increasingly concerned with notions of practical or successful intelligence (as distinct from academic of analytical intelligence) which he sees as involving the tacit knowledge that comes from experience and amounts to informed common sense. Thus Sternberg describes a social intelligence that involves social cognition as well as social judgement and practical intelligence (Sternberg 1990; Sternberg 1996).

The componential sub-theory involves the relation of intelligence to the internal world of the individual, in terms of the executive, performance and acquisition processes underlying task activities. This is the information processing aspect of the theory, concerned with meta-cognition and executive processes. Sternberg sees these aspects of intelligent behaviour as more or less invariant and universal.

Sternberg describes three components of cognition:

- metacomponents;
- performance components; and
- knowledge acquisition components.

He has identified the following metacomponents that organise and control information processing:

- deciding what the problem is;
- selecting;
- lower order components;
- forms of representation;
- strategies for combining lower order components;
- allocating attentional resources;
- solution monitoring; and
- evaluating external feedback.

He has identified the following performance components involved in task execution:

- encoding;
- inference;
- mapping;

- application;
- comparison;
- justification; and
- response.

Sternberg has also identified knowledge acquisition components that function to selectively encode, combine and compare information so as to facilitate the learning and structuring of new information. These functions involve:

- selective encoding;
- combination;
- comparison; and
- successive steps in task performance.

For Sternberg, intelligent behaviour entails the interaction of the 3 sub-theories of the triarchic model, but the notion of a general or overall performance remains important in his theory.

Sternberg's Critique of Conventional Intelligence Tests

Sternberg describes conventional intelligence tests as primarily analytic and unable to tap the creative and the practical forms of intelligence. He has contrasted the artificial, structured and closed nature of academic/analytic problems with the real-life, unstructured and open nature of practical problems. He has also contrasted the creative act of problem finding with academic problem solving. He has contrasted the characteristics that differentiate academic and practical kinds of problems in the following terms:

Table 12 Wagner and Sternberg's view of practical problem solving

Academic problems	Practical problems
Well defined	Ill-defined
Formulated by others	Unformulated
Necessary information provided	Additional information required
One correct answer	Multiple 'correct' answers
One method to obtain answer	Multiple methods to obtain answers
Disembedded from everyday experience	Embedded in everyday experience

(Adapted from Neisser (1976) and Wagner and Sternberg (1985))

Sternberg's Notion of Successful Intelligence

Sternberg has given increasing emphasis to what he calls 'successful intelligence' and the related recognition he gives to tacit or practical knowledge. Sternberg has increasingly emphasised practical problem solving and successful intelligence' in contrast with 'academic intelligence' He also contrasts general, academic and everyday intelligence. The successful intelligence described by Sternberg is related to lay views of intelligence in the emphasis given to practical problem solving, verbal ability, and social competence. Although Sternberg recognises that the psychometric notion of IQ has a certain meaning, he challenges the significance attributed to it by advocates to IQ (Sternberg 1996):

> IQ is predictive. What they fail to highlight, though, is that the prediction is quite weak (p.1).

He contrasts the importance of academic intelligence with that of successful intelligence:

> Academic intelligence of the kind measured by IQ tests matters, but really, it doesn't matter that much. For success in school and later in life, successful intelligence is key (p.4).

He distinguishes academic from successful intelligence:

> In my own work, I have drawn a rather sharp distinction between academic intelligence, on the one hand, and successful intelligence, on the other (Sternberg, 1996). Successful intelligence is the kind of intelligence you need to succeed in the real world. It is the translation of underlying skills and abilities into routines that lead to highly competent, everyday performances-on the job, in personal relationships, and in other aspects of our daily lives (p.1).

Sternberg describes successful intelligence in terms of tacit knowledge and skill:

> Successful intelligence is the acquisition and use of what you need to know to succeed in a particular environment, which you are not explicitly taught and which usually isn't even verbalized. Ask yourself how many aspects of your job are things you learned how to do in college or graduate school, and how many are things you learned on the job. Ask yourself how many are things someone showed you how to do, and how many are things you figured out for yourself, probably without anyone saying much of anything. (p.2).

Sternberg claims to have established the empirical substance of the difference between academic and successful intelligence:

> First, successful intelligence is distinct from academic intelligence. In study after study, the correlation between conventional measures and our measures is zero or close to it. In other words, you cannot predict a person's successful intelligence from his or her academic intelligence, or vice versa (p.3).

The Relationship of Sternberg's Theory to Other Theories of Cognition

Sternberg is concerned with those cognitive processes and capacities that can be applied to all kinds of content in all kinds of domains. He does not subscribe to the view that cognitive performance is context or domain-specific or that there is little that can be usefully generalised across contexts or domains. Ceci compares Sternberg's theory with that of other transdomainal theories in the following terms:

> According to the triarchic theory, executive components such as planning and revising are important in virtually all meaningful endeavours, and the existence of ***g*** is a testimonial to their wide-spread use in all domains.
>
> Sternberg departs from Eysenck, Jensen, and others in this view because he does not agree with them that the basis of variation in individual difference is at the level of microprocesses such as encoding and memory scanning. So, while they are inherently ***g***-based theorists, he is not. Indeed, for Sternberg, the source of the variation in individual differences is at a higher level of analysis, namely at the executive level or metalevel. This is probably why his theory is so attractive to educators: intelligence, according to his account, is congenial to remediation in that theories based on microlevel differences (e.g., encoding efficiency) are not. It is conceivable that someone may be trained to plan, revise, and replan; it is not clear that a person may be trained to use his or her microlevel processes more efficiently (Ceci 1990).

Although Sternberg gives recognition to the importance of context and experience in intellectual performance, there is still an emphasis in his work on general cognitive processes that function across contexts and classes of tasks. Although Sternberg has identified different kinds of intelligence, he does not accept the kind of pluralisation of the term intelligence seen in the work of Gardner. He has guardedly criticised Gardner's multiple intelligences theory as muddying the usual and useful distinction between general intelligence and special talents (Sternberg 1990).

A Response to Sternberg

A major virtue of the work of Sternberg is its comprehensiveness and eclecticism. He has, of course, a great deal of successful intelligence himself, and he does not judge it to be a successful strategy to rule anything out of consideration. In debates that can be highly politicised, Sternberg is a balanced and fair-minded commentator.

However the componential analysis which was the basis of his early work has not been particularly powerful, informative or fruitful. It is rather more analytical than empirical, and it has not really analysed higher cognitive processes into basic processes. As a result, it has not bridged the gap between psychology and biology.

It might be argued that Sternberg has been less focussed and rigorous than he might or should have been in his various researches. He is prepared, for instance, to use comparatively unanalysed concepts like practical intelligence, and they are not subject in his work to the kind of psychometric or psycho-biological exploration that should be the goal of cognitive psychology. Nevertheless, Sternberg's triarchic theory offers the best available overview of cognition. Although it is important for its comprehensiveness, the triarchic theory does not offer any really decisive insights, and it could have been more comprehensive, as will be shown later. While Sternberg is an important and influential critic of conventional intelligence testing, his triarchic theory has not changed the testing of intelligence, and he has not been able to change the dominance of conventional intelligence tests as he would have hoped (Sternberg 1991).

Psycho-biological Theories

Howard Gardner

The Stanford psychologist Howard Gardner has presented an influential theory of differential ability under the heading of multiple intelligences (MI) (Gardner 1993). Gardner is primarily a developmental psychologist. His work was initially inspired by Piaget, but it also grows out of neuro-psychology (Gardner 1974). His work also shows a strong interest in the expression of intelligence in the arts (Gardner 1973). In his early work at the Veterans Administration Medical Centre, Gardner studied brain damage and the breakdown of cognitive functioning so as to explore the neurophysiological basis of performance as a way of understanding the structure and organisation of symbolic functioning. This work involved studies of the localisation of brain function and out of it he developed his own differential ability theory in MI (Gardner 1985).

Gardner aimed to expand the scope of cognitive and developmental psychology towards biological and evolutionary roots of cognition, and in another direction towards cultural variations in cognitive competence. He sought the biological bases of cognition and also sought to accommodate cross-cultural perspectives in an understanding of intelligence.

Gardner explicitly rejected a unitary view of intelligence and he set out to find an 'empirically grounded set of faculties'. He eschewed apriori distinctions between types of knowledge or skill, such as Hirst's seven forms of knowledge (mathematics, physical sciences, interpersonal understandings, religion, literature and fine arts, morals and philosophy) or Gross' five modes of communication (Lexical, socio-gestural, iconic, logico-mathematical, musical) (p.59). He opposed the notion of general intelligence or ***g*** and he

also showed no interest in (or much knowledge of) the psychometric tradition, in his early work.

Gardner contrasted his view of multiple intelligences with differential ability theories developed by factor analysis because these theories:

- do not question the existence of general horizontal abilities;
- do not make intellectual contact with biology; and
- only sample a range of performances and are limited to pencil and paper tests.

According to Gardner, the psychometric view 'provides only that extremely partial glimpse of intelligence that reflects the Western scientific ethos' (Gardner 1985).

Gardener's Definition of an Intelligence

Gardner based his search for kinds of intelligence on a set of biological and psychological specifications, as well as on the notion that an intelligence must capture a reasonable gamut of the kinds of abilities valued by human cultures.

According to Gardner:

> An intelligence is an ability to solve problems, or to create products, that are valued within one or more cultural settings - a definition that says nothing about either the sources of these abilities or the proper means of 'testing' them (Gardner 1985).

The emphasis in this definition is on activity (problems and products) but it is extended (or qualified) by the notion of being an intelligence being valued in one or more cultural settings. This definition is culturally relativistic in stipulating that an intelligence has to be valued in a cultural setting. Eschewing the issue of testing intelligence in this definition is insignificant, however, eschewing any specification of the 'source' of the abilities is curious. The strength of Gardner's argument is in the psycho-biological grounding of his view, but his definition does not incorporate this characteristic. The definition does not require that an intelligence be grounded either in cognitive processes or structures.

Gardner's MI theory is based on an opposition to general or 'horizontal' powers:

> I strongly challenge the notion of large general powers ...the mind has the power to deal with several different kinds of content, but a person's facility with one content has little predictive power about his or her facility with other kinds. ...performance is likely to be specific to particular contents: human beings have evolved to exhibit several intelligences and not to draw variously on one flexible intelligence (p.IV).

He makes his challenge to a single unitary intelligence explicit:

> I seek to replace the current, largely discredited notion of intelligence as a single inherited trait (or set of traits) which can be reliably assessed through an hour-long interview or a pencil and paper test (p.286).

As used by Gardner, the term intelligence could be replaced with terms such as 'intellectual competencies', 'thought processes', 'cognitive capacities', 'cognitive skills', 'forms of knowledge':

> What is crucial is not the label but, rather, the conception: that individuals have a number of domains of potential intellectual competence (Gardner 1993).

According to Gardner, an intelligence is a relatively independent module with a core information processing mechanism that operates according to its own principles. These modules can be activated without central executive control, and they can be fashioned and combined in a multiplicity of adaptive ways by individuals and cultures. Gardner has a certain sympathy with the strong nativist arguments of Noam Chomsky (Chomsky 1986) and the related arguments of Jerry Fodor (Fodor 1983) about the modular basis of cognition.

Gardner sees a psycho-biological basis for each intelligence:

> I believe that at the core of each intelligence there exists a computational capacity, or information processing device, which is unique to that particular intelligence, and upon which are based the complex realisations and embodiments of that intelligence (Gardner 1985)

According to Gardner, these independent modules have their own perceptual and memory resources.

Gardner's Criteria for Defining an Intelligence

Gardner approached his exploration of cognition on the basis of 8 signs or criteria for identifying an intelligence. These criteria give a biological and psychological specifications for an intelligence:

1. breakdown by brain damage or pathology;
2. prodigies and idiot savants or selective absence;
3. core information processing operations;
4. distinctive developmental history;
5. definable expert end-state;
6. evolutionary history;
7. support from experimental psychology; and
8. encoding in a symbol system.

Although Gardner gave particular emphasis to the biological basis of intelligent performance, he also argued that an intelligence 'must be useful and important at least in certain cultural settings.' On the basis of the criterion of cultural significance, Gardner excluded gustatory and olfactory senses as intelligences, but he claimed that social roles in themselves do not qualify as intelligences (p61).

Clearly the major emphasis in Gardner's 8 criteria is on psycho-biology, and 5 of the 8 criteria are explicitly psycho-biological:

- breakdown by brain damage or pathology;
- prodigies and idiot savants or selective absence;
- distinctive developmental history;
- evolutionary history; and
- core information processing operations.

The criteria developed from Gardner's initial interest in brain damage, selective absence and developmental psychology. In Frames of Mind he extended these interests into evolutionary history, information processing and symbol systems. The emphasis on symbolic systems (the Vygotskian influence on Gardner's work) grew during the 1980s.

In building his theory, Gardner refers to evidence from experimental psychology, but most of the evidence for this theory is from social psychology (particularly in respect to the personal intelligences), and he makes little reference to (and shows little knowledge of) the substance of psychometric exploration of abilities. Significantly, Gardner does not compare his intelligences with the differential ability theories of Thurstone, Vernon, Guilford or Cattell outlined earlier in this chapter.

The most problematic aspect of Gardner's criteria for the definition of an intelligence is the 'definable expert end-state'. This pushes the notion of an intelligence in the direction of a social definition, that is a definition based on what a society sees as an important role. In one sense this seems common sense: why consider, let alone designate, some notion as an intelligence if it is not identified as important by a culture in general and is not manifested in important and valued roles? Such view is common sense rather than scientific view, and it is potentially at odds with the essentially psycho-biological emphasis that is the strength of Gardner's approach.

In Frames of Mind Gardner reviewed the evidence available on each of his 8 criteria to argue for the existence of an intelligence. He described the process as a kind of 'subjective factor analysis'. While he does not explain this comment, it seems that what he takes from factor analysis as a kind of activity is the aim to identify a small number of abilities that account for as much performance as possible with as little overlap or redundancy as possible. Reducing complexity to a small number of variables is, as we have seen above, the aim of factor analysis.

Gardner's Critique of Other Approaches

The research reported in Frames of Mind involved substantial critique of other notions of intelligence. Gardner criticises the psychometric, information processing and Piagetian approaches to intelligence for:

- focusing on logical and linguistic problems;
- not dealing with higher levels of cognitive functioning and creativity;
- ignoring the biological basis of cognition; and
- being insensitive to the range of roles highlighted in human society (p.24).

He stigmatises the 'IQ movement' as blindly empirical and concerned only with the predictive power of cognitive measures. According to Gardner, this approach offers no view of cognitive processes and no theory of how the mind works. It is a microscopic or shotgun approach. The test instruments used in such an approach are heavily dependent on language knowledge, biased towards crystallised intelligence and remote from everyday thinking. According to Gardner, the positive manifold at the heart of psychometric ***g*** is a result of all measures being pencil and paper tests that rely heavily on linguistic and logico-mathematical intelligences.

Gardner distinguishes his approach from the multi-factorial views within the psychometric tradition in that the psychometric approach does not:

- sample an adequate range of performances in an adequate range of assessment modes;
- make contact with the biological sub-strate of performance, or
- produce an adequate or comprehensive view of cognition (p.322).

Although he expresses admiration for the subtle penetration of Piaget's analysis of cognitive development, Gardner criticises Piaget's monolithic emphasis on a certain kind of thinking and one sort of development. According to Gardner, Piaget believed that there are general cognitive operations that can be applied to any content and that the development of these operations takes place in an invariant sequence of stages.

Gardner reports research into development as having shown that, rather than being a series of stages, development is a matter of continuous and gradual changes. Gardner describes research as showing considerable unevenness in the understanding and use of various concepts across domains. The attempt by Piagetians to accommodate this unevenness in the concept of 'decalage' is described as a 'fudge factor' by Gardner (p.21). According to Gardner, variation of development across domains is the rule rather than the exception.

Although Gardner admires Piaget's complex tasks and the experimental ingenuity with which he explores them, Gardner describes the matters explored by Piaget as remote from the kind of thinking most individuals undertake in their daily lives. According to Gardner, the work of Piaget is limited by its assumption of the role of the scientist as the goal of development, so that Piaget's view does not accommodate diverse roles and domains.

In Gardner's view the Piagetian enterprise fails to define a universal pattern of cognitive growth which all normal children are claimed to traverse, and he describes Piaget's work as restricted to the exercises of the science class:

> The classical structuralist view of intellect is even narrower than that held by the intelligence-test makers and in fact, is restricted almost completely to logical-mathematical thinking (p.323).

Gardner criticises the information processing approach to intelligence because it is also based on hypothesising a single horizontal problem solving ability. Gardner sees a computer-based view of thinking as excessively mechanistic and as only directed to, or dealing with, scientifically oriented problems. Such an approach is concerned with the microstructure of tasks and shows little interest in open-ended creativity, in Gardner's view. He also criticises the information processing approach for not offering a biological perspective on the nature of tasks that are employed in such research and because it lacks a theory about what constitutes the domain under scrutiny. According to Gardner, the exploration of information processing is an approach rather than a theory because it does not offer a coherent picture of human intellectual abilities (p.324).

Table 13 Gardner's Multiple Intelligences

The Seven Intelligences as defined by H. Gardner		
Intelligence	**End-States**	**Core Components**
Logical and numerical	Scientist Mathematician	Sensitivity to, and capacity to discern logical patterns; ability to handle long chains of reasoning.
Linguistic	Poet Journalist	Sensitivity to sounds, rhythms, and meanings of words; sensitivity to the different functions of language.
Musical	Composer Violinist	Abilities to produce and appreciate rhythm, pitch and timbre; appreciation of the forms of musical expressiveness.
Spatial	Navigator Sculptor	Capacities to perceive the visual- spatial world accurately and to perform transformations in one's initial perceptions.
Bodily-kinaesthetic	Dancer Athlete	Abilities to control one's body movements and to handle objects skilfully
Interpersonal	Therapist Salesperson	Capacities to discern and respond appropriately to the moods, temperaments, motivations and desires of other people.
Intrapersonal	Person with detailed, accurate self-knowledge	Access to one's own feelings and the ability to discriminate among them and draw upon them to guide behaviour; knowledge of one's own strengths, weaknesses, desires and intelligences.

Gardner sees the use of symbol systems as a distinctive element of human cognition, and he sees human information processing as the deployment of various symbol systems. For Gardner, the ability to discern different symbol systems is an important element of understanding intelligence. While he seeks to use the methods and overall scheme of Piaget, he focuses them on more than linguistic, logical and numerical symbols. For Gardner, the challenge is to compose a developmental portrait of each of these forms of symbolic competence and to determine empirically which connections or distinctions might apply across them. In this way, Gardner seeks to reconcile a pluralist approach to cognition with Piaget's unilinear developmental scheme.

Gardner does not accept the notion of defining intelligence on logical grounds. He recognises the strength of a neuro-psychological approach based on the exploration of damage to, and breakdown of, brain function. He also recognises the criticism that the physical proximity of various activities and functions in the nervous system may not necessarily reflect modular organisation of functions. Gardner gives increased emphasis in the second edition of Frames of Mind to the view that domains are what are designated by a culture, and that cognition develops through socio-cultural interactions.

Gardner aspires to synthesise or reconcile nativist, epistemic, developmental and the contextualist views with research into the biological foundations of intelligence:

> Somewhere between the Chomskian stress on individuals with their separate unfolding mental faculties, the Piagetian view of the developing organism passing through a uniform sequence of stages, and the anthropological attention to the formative effects of the cultural environment, it ought to be possible to forge a productive middle ground: a position that takes seriously the nature of innate intellectual proclivities, the heterogeneous processes of development in the child, and the ways in which these are shaped and transformed by particular practices and values of culture (p.328)

Gardner reviews research into cognitive development in Frames of Mind and concludes that, while the preponderance of evidence supports the plasticity of the brain in the early months of life, development is also modulated by genetic constraints. According to Gardner, while there is flexibility in human development that is shaped by culture, humans are predisposed to carry out certain specific intellectual operations.

Gardner is drawn to biological approaches to cognition because he hopes that 'specificity of cognitive function can be tied to much finer regions of the human cerebral cortex' (p51). He reviews the arguments advocating the equipotential or uncommitted nature of cerebral capacity, and he concluded in 1983 that the 'biological basis for specialized intelligences' (p.57) defining 'natural kinds of human intelligence' is almost achieved:

> The once seemingly unbridgeable gap between behaviour and biology seems to have been sealed (p.47).

Gardner recognises the influence of genetic inheritance (genotype) but he also recognises the importance of the way characteristics are expressed in a given environment (phenotype) so that he is sceptical about the notion of specific heritability of intelligence:

> There is considerable agreement that physical traits are most straight forwardly genetic, that aspects of temperaments are also largely genetic, but when it comes down to aspects of cognitive style or personality, the case for high heritability is far less convincing (p.34).

Gardner's Intelligences

Although Gardner recognises the amazing fact that the brain is sufficiently plastic (equipotential) in early life to allow removal of the left hemisphere but still achieve the development of normal language function, he concludes that language is 'a picture of astonishing specificity and localisation in the brain' (p.87).

Gardner recognises that strength of Chomsky's argument about a genetically hard-wired language acquisition device (Chomsky 1980), but he also recognises the influence of culture

on language development in that 'canalisation' governs processes of language acquisition. Gardner concludes his review of the biological basis for a linguistic intelligence by stating:

> In terms of my 'criteria' we might say that syntax and phonology lie close to the core of linguistic intelligence while semantics and pragmatics include inputs from other intelligences (p.80).

However this comment is close to recognising that although the basic components of syntax and phonology are specifically localised, the more complex functions of semantics and pragmatics are not localised:

> It is difficult to find a fully fledged compromise of one's language area in the face of otherwise unimpaired comprehension and reasoning skills (p.88).

Thus even the most specifically located of Gardner's intelligences also involves many other aspects of the brain, particularly in the general functions of interpretive thought.

Gardner describes logico-mathematical thought as developing from the confrontation of the individual with the world of objects. He sees Piaget as offering a brilliant pattern of development in this one domain, even though Piaget's stages do not apply to all domains of development.

Gardner describes the developmental trajectory of logico-mathematical thought as less regular, lockstep and stage-like than Piaget would have wished. According to Gardner, rather than moving through definite stages, development proves to be gradual and heterogeneous. After reviewing the issues of brain localisation of mathematical/logical functioning, Gardner concludes that neurological organisation underlying such functioning is less localised than those for language or music:

> Unlike logico-mathematical thinking, language and music are comparatively robust in the face of general intellectual decline provided specific areas have not been singled out (p158).

Gardner owns an attraction to the notion that logico-mathematical thinking may not be a relatively autonomous system, and even contemplates the notion that it is a supra- or general intelligence:

> Wielding Ockham's razor, one could conclude that logical-mathematical ability is not a 'pure' or 'autonomous' a system as others reviewed here, and perhaps should count not as a single intelligence but as a kind of supra- or more general intelligence. I have at times had sympathy for this argument and do not want, in these pages, to come off as more definitive than I actually feel. However, in my view, the fact that one can encounter specific and particular breakdowns of logico-mathematical ability, as well as examples of extreme precocity, makes the elimination of logical-mathematical intellect far too extreme a scientific manoeuvre (p.159).

Although Gardner feels constrained to consider whether the logico-mathematical intelligence is 'more basic than others' or 'at the centre of human intellect' (p168), these doubts do not undermine his scepticism about a general intelligence.

Gardner's review of the neurological basis of spatial intelligence recognises a propensity for the involvement of the right hemisphere in spatial tasks, but he concludes that the neural basis of spatial intelligence is not firmly established (p.240).

Gardner's advocacy of a pair of personal intelligences is not, in itself, particularly unconventional (Sternberg 2000), but the comparative attention he gives to developing the notion is significant. As was noted earlier in this chapter, Spearman gave particular emphasis to the notion of social cognition, as did Guilford and David Weschler, the creator of the influential WISC intelligence tests. But Gardner is justified in claiming (in 1983 at least) that such thinking about personal intelligences has 'tended to be ignored or minimized by nearly all students of cognition' (p.242). Through the work of Mayer and Salovey (Mayer 1989) on emotional intelligence, and the popularisation of it by Daniel Goleman (Goleman 1995), and the work of Antonio Damasio (Damasio 1994; Damasio 1999) and Joseph LeDoux (LeDoux 1996) on the relationship between emotion and reason, social cognition has received significant attention in the 1990s.

Gardner contends for two kinds of personal intelligence:

- intra-personal, involving access to one's own feelings; and
- inter-personal, involving the ability to notice and make distinctions about the feelings of others.

According to Gardner, these are the most encultured of intelligences. They can take various forms, and they have a wide range of end-states, but he views these intelligences as of 'tremendous importance' (p.242).

According to Gardner, there is evidence for the right hemisphere involvement in emotional, spatial and interpersonal thinking:

> Defects in the frontal lobe can interfere with the development of personal forms of knowledge and can cause various pathological forms of intrapersonal and interpersonal knowledge (p.262).

He concludes that on the original list of criteria personal intelligences 'pass muster rather well' (p.244) as they have:

- an indentifiable core;
- a characteristic pattern of development;

- specifiable end-states;
- neurological representation and a pattern of breakdown; and
- symbolic systems in rituals, myths, religious codes.

Gardner even speculates about the involvement of the personal intelligences in that most fundamental characteristic of an individual, the 'sense of self'. Recent evidence, particularly the work of Antonio Damasio, has given strong support to the neurological basis of social cognition and personal intelligences (Damasio 1999).

A Response to Gardner

Gardner is an excellent synthesiser of information and he produces interesting and readable books. The Shattered Mind is a model of technical writing in an accessible form, and The Minds New Science is a very interesting and informative overview of the rise of cognitive science. Gardner likes a good story and he can tell one very well. In The Minds New Science, for instance, he presents the drama of the young Chomsky's challenge to the behaviourist establishment in a few deft touches. In some respects The Arts and Human Development is the most ambitious of Gardner's works in its attempt at a theory of development in the arts, but it lacks clarity and convincingness. Gardner's strength is narrative synthesis rather than theory.

Gardner has also been very ambitious but not particularly lucky in his predictions. In The Quest for Mind, for instance, he heralded the importance of structuralism just as post-structuralism and deconstruction were gathering momentum. Similarly, Gardner celebrated the imminent bridging of the gap between psychology and biology in 1983, but the issue has turned out to be much more complicated and equivocal than he envisaged. There are few who would claim at the beginning of the C21 that the bridging of the gap between psychology and biology is imminent (Greenfield 1997; Calvin 2001; OECD 2002).

Gardner is an intelligent man, and Frames of Mind is an interesting and readable book, but it is not rigorous and strong. It is not clear or coherent about some basic issues. He is equivocal, for instance, about the extent to which his intelligences are biological structures or social constructs. Frames of Mind does not deal with all of the views and arguments it should have, and some of the material presented is merely gestural. The later text book, Intelligence: Multiple Perspectives (Gardner, Kornharber et al. 1995), which Gardner co-authored in 1995 is a more thorough and comprehensive examination of the basic issues than Frames of Mind. It seems that Gardner learned some of the things he should have covered in developing Frames of Mind a decade or so after having developed his MI theory.

Frames of Mind skates over an enormous amount of territory but it is not powerfully focussed or convincing. It seems to be a book that arose from the sort of political passion that prompted Gould's The Mismeasure of Man, and it seems to take up a position rather than examining the evidence and arguments to draw a conclusion.

It is reasonable for Messick to identify under the heading 'Ideology and Evidence' a political impulse behind Gardner's writing on intelligence:

> One reason for emphasising full autonomy is an ideological commitment implicit in Gardner's writings that independence of functioning signals equal status or importance for a variety of ability domains, thereby freeing education and human development from the confines of general intelligence. Key questions are whether semi-autonomous modular functioning would serve well enough for those purposes and what Gardner might sacrifice in educational practice by downplaying the operation (and training) of general cross-cutting processes such as memory, reasoning and judgement (Messick 1992).

One can sympathise with the egalitarian impulse behind Gardner's argument for multiple intelligences, but one can also recognise that Gardner, like Gould, treats the issue of intelligence so that the political passion overpowers the science. While theories of how the mind/brain functions have often been related to various social and political views, attempts to theorise the mind/brain and thought processes should aspire to be scientific rather than political. It is arguable that Gardner's MI is rather more political that it is scientific.

It is fair to say that Gardner simply expands the term 'intelligence', and it is arguable that in doing so he confuses and debases it. It was usual for psychologists to use the term 'intelligence' to refer to the notion of a general cognitive ability and to use the terms 'special', 'broad' or 'specific' abilities for different elements of a differential theory of cognitive ability. It is this useful distinction that Gardner collapses.

When one looks at the substance of Gardner's theory one sees that there is nothing surprising about the linguistic, logico-mathematical, spatial and personal intelligences he identifies, and these are clear similarities between Gardner's intelligences and the abilities of the psychometric tradition outlined above. It is in the bodily/kinaesthetic and the musical intelligences that Gardner collapses the basic distinctions between cognition, affect, conation and psychomotor activity. Gardner's designation of musical activity as an intelligence loses the distinctively affective element of music.

One may place great value on various bodily performances or musical activities without having to claim that they are more or less the same as more cognitive activities. The distinctive feature of more cognitive activity is that it is internally cognitive rather than

psychomotor. And the distinctive feature of music is that it is a kind of thinking that is suffused with emotion, and perhaps predominantly emotional.

Basing his argument on the biological substrate of cognition is a strength of Gardner's approach, but the neurological evidence is less powerful than Gardner hopes, and it does not decisively support meaningful modularity of mind as the basis of complex performances.

Gardner muddies thinking about cognition by introducing the notion of an 'end-state' as a necessary requirement for an intelligence. This effectively conflates the psycho-biological consideration with social roles. Although the notion of end-states is important in Gardner's definition of an intelligence, with time he has had to recognise the diverse abilities involved in any end-state and to performance in any domain. Gardner wants to identify his intelligences with certain social roles but different roles involve all sorts of abilities. He gives little attention to defining the basis for different kinds of end-states and he treats this crucial issue as more or less self-evident. In the end, the notion of an end-state does not turn out to define a cognitive capacity.

After he developed his MI theory, Gardner was confident that it had specific and practical meaning in education and educational assessment (Gardner 1991; Gardner 1999). One suspects that MI theory would prompt a valuable but rather vague catholicity in pedagogy, but there is reason for thinking that MI theory has little to offer assessment, particularly at the hard edge of summative assessment and certification (Gardner 1992). It seems that Gardner discovered his MI theory did not have much meaning or use in assessment as he came to recognise that different aspects of his intelligences are involved in all sorts of performances (Gardner 1993): 'Intelligences typically work in harmony so their autonomy may be invisible' (p.9).

If intelligences 'typically work in harmony so their autonomy may be invisible', one can only wonder what the notion that they are autonomous means. Gardner even comes to believe that it is doubtful whether it is possible to measure the core ability of an intelligence, at least in a way that is 'knowledge free'. He concludes that

> A central implication of MI theory, then, is the need for 'intelligence fair' assessments which allow participants to engage in real-world activities and use relevant background knowledge to solve problems. (Gardner 1993).

The desire to find an 'intelligence fair' assessment was an important implication of MI theory, but the notion of the 'core intelligences' seems to evaporate in 'real-world activities'. Eventually it seems that Gardner sees assessment as a matter of 'real world activities', but he does not offer a taxonomy of activities or a basis for defining types of activity.

The notion of a verbal or linguistic intelligence is fundamental in the psychometric tradition. The 'end-state' of the linguistic intelligence for Gardner is the poet and the journalist, which seem insignificant in comparison with the scientist and the mathematician of the logico-mathematical intelligence. Gardner's linguistic intelligence is a rather poor cousin of the psychometric concept of verbal ability. Gardner rightly objects to the Piagetian notion of the scientist as the goal of development, but his classification gives more to this hegemony than does conventional psychometrics.

The logical intelligence nominated by Gardner is also quite conventional, but turning it into mathematics and science is to turn it into a domain (or perhaps a sub-set of crystallised abilities) rather than an ability. Again, the impact of Gardner's conflation of cognitive capacities with social roles turns out to be a narrowing of the psychometric concept of abstract, formal and logical reasoning.

After publishing Frames of Mind, it seemed to Gardner that the pluralising of intelligence was going too far, particularly with the personal intelligences. He is uncomfortable with the notion of emotional intelligence developed by Mayer and Salovey (Gardner 1993). He has recently attempted to limit the scope of intelligences and sought to clearly distinguish them from personality and from issues of values and ethics:

> Note, however, that MI theory makes no claims whatsoever to deal with issues beyond the intellect. MI theory is not, and does not pretend to be, about personality, will, morality, attention, motivation, and other psychological constructs. Note as well that MI theory is not connected to any set of morals or values. An intelligence can be put to an ethical or an antisocial use. Poet and playwright Johann Wolfgang von Goethe and Nazi propagandist Joseph Goebbels were both masters of the German language, but how different were the uses to which they put their talents!

The personal intelligences have been commonly identified, although they have not been explored as much as they deserve. Such exploration is currently taking place under the heading of emotional intelligence, but, ironically, Gardner challenges this notion. He argues that he is concerned with cognition and he implies that a notion of emotional intelligence collapses the important distinction between emotion and intelligence. One can only wonder why collapsing the distinction between cognition and character is any more objectionable than collapsing, as Gardner does, the distinction between cognition and movement.

Sternberg (Sternberg 1989) criticises Gardner for collapsing the distinction between cognitive abilities and talents:

> ... an ability is a component of intelligence when we cannot get along without it and a talent when we are not noticeably handicapped by it absence (p.42).

For Sternberg talents are specialised and an intelligence is general.

Although MI has gained a general popularity it lacks empirical grounding, it has not been empirically fertile, and it has not been subject to much empirical testing (Kearsley 1994).

Gardner's collapsing of the distinction between cognitive abilities and musical and kinaesthetic abilities is innocuous. Giving greater recognition to musical and kinaesthetic abilities is a good thing, and there is no danger that these kinds of performance will be taken to be basic abilities or compulsory requirements. We are not in danger of compelling students to study music or to dance or to play sport. We do, on the other hand, require basic levels of literacy and numeracy, and advancement in the education system depends on literacy and other cognitive learning. The way in which we conceptualise and partition cognition can have a significant impact on what we look for and do in education.

The dangerous aspect of Gardner's work is his conflation of abilities with social roles. This conflation is a problem because it effectively dissolves the notion and the definition of cognitive abilities into or within social roles. What is really important in Gardner's conceptualisation are the notions of the:

- scientist and mathematician;
- poet and journalist;
- composer and violinist;
- navigator and sculptor;
- dancer and athlete; and
- therapist and salesperson.

The equation of abilities with social roles in this fashion is to dissolve the notion of generic and underpinning abilities on the presumption that there is a fundamental difference in the cognition undertaken by the scientist and the poet.

Gardner's Responses to Criticisms of MI Theory

In the 1993 revision of Frames of Mind Gardner offered a significantly more sophisticated version of the relationship between biological and the socio-cultural elements of intelligence than in 1983. In the later edition he described intelligences as 'intellectual proclivities' that 'may be thought of in neurobiological terms'. Such a description is to effectively disclaim the psycho-biological basis for his intelligences. He also came to recognise the importance of domains and fields on what counts as an intelligence. According to Gardner, a domain is an area of knowledge housed in cultures that amounts to impersonal expertise, whereas a field is a social structure that determines the basis of and recognises competence. It seems

that Gardner is forced to make these distinctions in response to criticisms that note a certain incongruity in the biological and socio-cultural aspects of his criteria for an intelligence.

By 1993 Gardener seems equivocal about the substance of his intelligences. He is concerned with the biological basis of intelligences, but in the second edition of Frames of Mind he says of the word 'ntelligence'that:

> ... we have come to believe in its existence, as genuine tangible, measurable entity, rather than as a convenient way of labelling some phenomena that may (but may well not) exist (p.69).

He recognises that there is a universal human temptation to reification of constructs, but he explicitly repudiates such reification for his own theory:

> These intelligences are fictions - at most useful fictions - for discussing processes and abilities that (like all life) are continuous with one another. ...they exist not as physically verifiable entities but only as potentially useful scientific constructs (p.69).

A statement like this by Gardner seems to accept that his intelligences are mere constructs that do not claim to be psycho-biological structures at all.

Jensen dismisses the claim that Gardner's intelligences are independent on the basis that there are well-established correlations between at least four of these 'intelligences' (verbal, logical-mathematical, spatial, musical), all of which are substantially ***g*** loaded (Jensen 1998). As the factorial structure of two of the personal intelligences has not been determined, their ***g*** loadings remain unknown in Jensen's view, and he challenges the notion of the kinaesthetic as a cognitive ability. Jensen, Messick (Messick 1992) and Carroll (Carroll 1993) all argue that there is no incompatibility between ***g*** and the existence of neural modules that control particular abilities. These claims challenge the fundamental aim of Gardner's argument.

Other theorists, particularly those under the influence of Vygotsky (Wertsch 1985), have dealt with the socio-cultural and the symbolic nature of intelligence more coherently and consistently than Gardner.

Symbol System Theories

L.V. Vygotsky

The discovery of the Russian psychologist, Lev Vygotsky (1896-1934), by the English speaking world in the 1960s and 1970s was both a result of and an influence on the growing emphasis on the socio-cultural aspects of cognitive ability. In his work, Vygotsky had been interested in explaining higher mental functions like concept development, problem solving

and language development, and he gave particular emphasis to the way higher mental functions in humans were based on, and involved the use of physical and symbolic tools.

According to Vygotsky, physical and symbolic tools are invented by a culture and are mastered by an individual through socialisation processes (Vygotsky 1986). Vygotsky emphasised the way higher mental functions are tied to social interactions, and the way cognitive development is largely influenced by a child's interactions with others. According to Vygotsky: 'All higher mental functions are internalised social relationships' (p.164).

Vygotsky differed with the Piagetian view that cognitive development was based on accommodation of the individual to the world. For Vygotsky, cognitive development was based on assimilation of a particular socio-cultural world by the individual.

Vygotsky was critical of decontextualised testing of cognitive performance. He argued that cognitive performance had to be understood within a cultural situation, and from this perspective a test should be understood as a dynamic social situation in itself. These views led to Vygotsky's distinction between a level of performance in a decontextualised test and the 'the zone of proximal development' of an individual on the same problems, given the support and assistance of someone more knowledgeable.

The notion of the 'zone of proximal development' has significant implications for assessment and learning. It leads to the notion of dynamic testing that involves offering information and ideas to a test taker in the course of problem solving, and it emphasises the importance of external guidance or 'scaffolding' as a means of learning. These ideas are being developed and extended by those who take a situated or contextualist view of cognition (Collier 1994).

Socio-cultural or Contextualist Theories

S. J. Ceci

Stephan Ceci of Cornell University drew on Sternberg's triarchic theory to develop what he called the bioecological (BE) theory of intelligence (Ceci 1990). While he shares something of Sternberg's emphasis on the biological components of cognition, he gives much greater emphasis to the contextual sub-theory than Sternberg and he pushes it much further. Ceci also emphatically differs with Sternberg's assumption about the transdomainal nature of cognitive processes.

Ceci shares the aims of Gould and Gardner in that he sets out to discredit what he calls the fallacy that 'biology = IQ = intelligence = singularity of mind = real-world success' (p.198). Unlike Gardner, Ceci does not ignore psychometric approaches to cognition, and unlike Gould he does not dismiss without consideration psychometric evidence about cognition.

Ceci's Response to the Psychometric Tradition and *g*

Ceci explicitly recognises what he calls the 'five easy facts of the intelligence research community' that:

- there is a positive manifold in test scores;
- factor analysis does give a first principal factor;
- many researchers accept the first principal factor as ***g***;
- there are positive correlations between ***g*** and academic and social accomplishments; and
- a number of studies report that ***g*** is quite heritable, with a correlation between the two of 0.5 - 0.6.

Unlike Gardner and Gould, Ceci accepts these facts of psychometrics, but he challenges in detail the nature, implications and significance of these facts.

Ceci does not accept that the positive manifold confirms the existence of ***g***, and he argues that there is no primary causal role for IQ in higher level understandings:

> A case can be made, and I will try to make it, that the relationship between IQ and advanced academic understanding is epiphenomenal; that persons with low IQs may have the cognitive architecture needed to appreciate advanced mathematics, language, and science, but they lack the relevant background experience. Were they provided such background experience, then both their IQs and their linguistic, scientific, and mathematical understanding might be elevated. IQ would be not the cause of the advanced understanding, but simply a concomitant attribute which is susceptible to the same type of experiential influences. It then becomes no more correct to assert that individuals are good

> at language, science, and mathematics because they are intelligent (i.e., have high IQs) than it is to assert that individuals have high IQs because they are good at mathematics, language, and science (p.8).

Ceci argues that intelligence is a multifaceted set of abilities and that a specific facet might become more or less effective as a result of the physical, social, cultural, and historical contexts in which it has been crystallized and in which it is subsequently assessed. According to Ceci:

> ...the generality of intellectual functioning is more illusory than real, a phenomenon inextricably linked to the paradigms, values, and assumptions made by those committed to a deterministic viewpoint (p.8).

Ceci challenges the significance of ***g*** on the basis of the importance of the impact of schooling on cognitive performance. He argues that IQ is academic performance and that it does not reflect non-academic abilities. Like Gardner, Ceci challenges the significance of the positive manifold because it is based on the narrow range of questions, materials and contexts in IQ tests, and hence, he claims, it is not surprising that such tests are inter-correlated.

Ceci's Bio-ecological Theory

Ceci describes his own BE theory as a complete person x process x context model:

> Readers may recognize that the foregoing depiction is a type of 'person x process x context' model. *Person* variables refer to biological bases for either cognitive or personality dispositions, while *process*es are mental operations, such as inferring, and *context* is broadly construed to refer to environmental potentiators as well as to the structure of one's knowledge. If the biological efficiency of a particular process is known, its actual efficiency would be a result of a series of reactions it had with contexts (p.4).

The bio-ecological theory posits that:

- intelligence is a function of the interactions between innate potential abilities, environmental context, and internal motivation;
- there are multiple cognitive potentials or intelligences;
- there are domain-specific processes;
- knowledge is fundamentally important to performance;
- knowledge and ability are interrelated;
- schooling has an impact on IQ;
- social class is an important determinant of IQ; and
- cognition is frequently context sensitive.

According to the BE theory, intelligence involves the interchange between biological and environmental resources, and performance involves both innate potentials and environmental resources.

Ceci draws data for his argument from a diverse range of disciplines (see Table 14) including: sociology, anthropology, history, genetics and education. He considers:

- micro level processing;
- macro level processing;
- the role of knowledge;
- reaction time data;
- the positive manifold; and
- life course of the processes.

Ceci looks for contextual influences on cognition rather than the 'usual emphasis on cognitive universals' (p.xiv). He takes a modular view of the mind, he seeks contextual determinants of intelligence, and he aims to go well beyond mildly cultural accounts, like that of Sternberg, of intelligent performance. Ceci defines the notion of context very broadly to include motivational forces, social and physical aspects of a setting or task, and values inculcated through various types of parenting. He claims that knowledge and ability are fundamentally inseparable.

Ceci's view of intelligence identifies basic biological processes and also recognises the importance of experience and knowledge. He recognises multiple forms of intelligence and his view is concerned with developmental processes. Rather than dismissing psychometric evidence, Ceci argues for multiple cognitive potentials on the basis of evidence from the psychometric tradition of differential abilities.

Ceci challenges Spearman's view of the importance of ***g*** and the claim made by Spearman about the 'indifference of the indicator' in measuring ***g***. Spearman believed that ***g*** would be measurable through any cognitive task whereas Ceci treats it as an unstable epiphenomenon. He attempts to demonstrate this experimentally by designing what he called 'problem isomorphs' that show different performances of the same skill in different contexts with different content. He contrasted the performance, for instance, of children on the same task in the form of an abstract laboratory task and in the form of a video game. The comparison of performance on these different tasks demonstrated that the abilities of the children changed significantly depending on whether the task was in the form of a laboratory task or a video game.

Ceci recognises the importance of domains of knowledge in cognitive performance, and he recognises the way knowledge and basic processes interact. Ceci sees knowledge and intelligence as inseparable:

> But it is illusory to imagine that performance on IQ tests or, for that matter, even microlevel cognitive tasks like the encoding of alpha-numeric stimuli, is devoid of knowledge. Nor can it be said that the detection of relationships can, even in principle, be estimated without reference to past learning and to the elaborateness of one's knowledge base (i.e., cognitive complexity). As Scribner (1986) has recently concluded on the basis of her analysis of everyday problem solving, the ability to solve problems is intimately tied to the amount and quality of relevant knowledge one possesses. She argues that 'From earlier assumptions that problem-solving can be understood in terms of 'pure process', a consensus has arisen that problem-solving procedures are bound up with the amount and organization of subject matter knowledge' (p.117).

Ceci described this interaction as 'cognitive complexity', a term he seems to use as a synonym for intelligence. This cognitive complexity is the extent to which one's cognitive processes can 'operate on one's knowledge structures in a complex, efficient, and flexible manner' (p.15):

> Cognitive complexity results from efficient processes operating in elaborated knowledge domains. It is what the lay person most likely means when marvelling at someone's skill or intelligence in a specific area (p.26).

In BE theory, knowledge and ability are inseparable:

> It is logically impossible to separate these intelligences from acquired knowledge (and the elaborateness with which this knowledge is organized), even though in principle such a distinction can (and in many theories must) be made (p.14).

Ceci summarises the relationship he sees between intelligence, cognitive complexity, knowledge and cognitive processes in the following fashion:

> ... intelligence is a function of cognitive complexity, which in turn is dependent upon the operation of cognitive processes on a specifiable knowledge structure and, conversely, cognitive processes are dependent upon the sheer quantity of knowledge a person possesses as well as the organization of this knowledge. And, finally, cognitive complexity is tied to specific domains of knowledge or information (p.27).

Ceci is a contextualist in that he believes the nature of intelligence is wholly or partly determined by what a culture values, and he argues for a relative independence of academic and practical kinds of intelligence. Unlike Sternberg, who focuses on the different kinds of intelligence involved in practical skills versus analytic thought, Ceci concludes that, at least in certain groups, intelligence tests just don't measure intelligence.

Ceci's Empirical Research

Ceci's empirical work dealt with the influence of context on thinking and reasoning, and he showed substantial changes in the way people solved problems as a function of their familiarity with the type of problem, its concreteness, and their level of motivation to solve it:

> It wasn't only that problem-solving performance could be elevated or lowered, depending on the context in which the task was performed (though that is an interesting finding in its own right), but sometimes individuals in these studies would behave in an exceedingly complex manner that was totally unpredictable from what we knew about them, including their IQ scores. Insofar as the mental complexity these subjects exhibited in solving these problems depended on aspects of the context as much if not more so than it did on their IQs, I began to question the interpretation of the latter as an indicant of disembedded, acontextual problem-solving aptitude (p.xiii).

Ceci tested the nature of expertise in wagering on harness races (Ceci and Liker 1986). He found that skilled handicappers implicitly used a highly complex interactive model with as many as seven variables in making predictions, and he also found that the ability to make such predictions successfully was unrelated to scores on intelligence tests. There was no correlation between handicappers' IQ and their handicapping performance.

Ceci draws on similar work by other researchers. He outlines the work of Carraher, Carraher, and Schliemann (Ceci 1990) in which it was found that Brazilian street children were capable of doing the complex mathematics required for survival in their street business even though they would have failed tests of school mathematics. Similarly, shoppers in California who had no difficulty in comparing product values at the supermarket were unable to carry out the same mathematical operations in paper-and pencil tests (Lave 1988).

Although BE theory is explicitly biological, and Ceci sees biological and componential evidence for intelligence as particularly important, he recognises that the evidence in not conclusive. In reviewing the kind of neurological results (such as that offered by Gardner) from patients with brain lesions for a modular structure of the brain/mind, he recognises that such evidence may indicate only that the specific functions that are impaired are subserved in part by specific neuro-anatomical systems - not that these functions are independent of some general neural factor:

> In short, that specific functions are lost as a result of localized brain insult is pretty clear; but that this implies the absence of a global biologically based cognitive factor (*g*) is not so clear, at least not until evidence is found that rules out the existence of some underlying cortical unit that performs a uniform operation it varies from region to region and area to area, according to extrinsic connections. So even when the neuro-psychological findings are used to argue in favor of the hypothesis of multiple cognitive potentials (and against the singularity of intelligence), they are problematic (p.105).

Ceci makes these qualifications particularly in relation to the biological basis of Gardner's theory, and he concludes that the biological evidence offered by Gardner is not convincing:

> A careful reading of Gardner's work indicates that the 'subjective factor analysis' that forms the basis for his seven intelligences was itself subjective, a claim I am sure he would be the first to acknowledge. (In his book, he described it as only a 'first cut' and subject to future revision.) But the best evidence he mustered for the existence of multiple independent intelligences was the neurological that showed that if one area of the brain was ablated, a particular function was lost. For Gardner, this above all else refuted the notion of ***g***. As was pointed out in an earlier chapter in which an analogy with athletic prowess was made, however, this neurological evidence is really not convincing and should *not* be seen as contradictory to the idea of ***g***. For that one must go beyond Gardner's arguments, while acknowledging the importance of his insights (p.210).

Although he presents a wide range of argument and evidence against what he sees as the fallacy of ***g***, Ceci's conclusion about the state of play in the theory of cognition is qualified and provisional:

> Recently, a number of experimental psychologists have begun to question the traditional notions surrounding intelligence and IQ, arguing for a more ecologically based conceptualization of cognitive complexity (Cole, 1975; Keating, 1984; Neisser, 1979). In the last four years, reports have appeared from experimental psychologists that support the anthropological findings previously cited. These studies demonstrate that even within a given cognitive domain, there are substantial mismatches between performance in one setting versus another. And across cognitive domains, there appears to be far less correlation between one measure of cognitive complexity and another than previously has been assumed to exist. Alone, none of these studies provide incontrovertible evidence for the contextual nature of intelligent behaviour, but together they call into question a prior era's unrivaled assumptions regarding the generality of intelligence and the mere adjunctive status of context. They demonstrate the importance of contextual variables as a *constituent* in the perception and solution of complex problems, rather than simply a sociophysical address at which cognition unfolds (p.36).

After amassing as much evidence and argument as he can to discredit the biological fallacy, this is a very balanced and circumscribed conclusion. Although he attempts to discredit the notion of transdomainal skills and argue for the importance of knowledge in skilled performance, Ceci does not offer a blanket dismissal of the notion of generic skills. He recognises that the argument between the veriticalists and the horizontalists is by no means concluded.

A Response to Ceci

In some respects the work of Ceci is more comprehensive than that of Sternberg or Gardner. He takes the experiential and contextual more seriously than Sternberg, and he deals with psychometric work that is ignored by Gardner. His arguments against ***g*** are significantly

more comprehensive and resourceful than those of Gardner, and they are more telling because they attempt to directly rebut the evidence offered in support of ***g***, rather than ignoring it.

However, the place of the biological in Ceci's theory is uncertain. He wants to ground his theory in biology, but he wants to deny that biology accounts for IQ and real-world success. He presents a persuasive range of arguments and evidence that differences in IQ are socially conditioned, but he seems mesmerised by ***g*** and he does not consider the notion that biology = multiple intelligences = real world success. Ceci would presumably describe such a proposition as a fallacy, but he does not refute it, and it is such a claim that makes much of his argument immaterial.

Ceci's argument is based on the claim that he can discredit ***g*** by showing that it does not account for the differences between ordinary and expert performers. But this argument does not discredit the claim of advocates of the significance of ***g***, that ***g*** predicts a certain facility in learning. Demonstrating that children perform differently on laboratory tasks and on games shows the importance of interest and motivation, but it does not invalidate the notion of transdomainal skills and the predictive power of ***g***. Expertise is a matter of rich knowledge structures, but ***g*** theorists do not claim to account for expertise. They wish to focus on the capacities that facilitate the development of expertise. They do not claim to describe the differences between novices and experts, rather they are concerned with the capacities that facilitate the development of expertise. Similarly, the fact that skilful performance of street children and shoppers is context-related does not invalidate the notion of transdomainal skills and the predictive power of ***g*** or various special abilities.

The fundamental claim made by Ceci, that we can never disentangle intelligence from knowledge, is not decisive because we can significantly vary the importance of knowledge in particular performances. Thus an argument, as entertained by Sternberg, that claims to focus on general intelligence by minimising the importance of particular knowledge, is strong.

Ceci is prone to confuse expertise with intelligence. While expertise is fundamentally a matter of rich knowledge structures, intelligence is best thought of as a matter of the flexible capacities that facilitate learning. To focus on intelligence is to focus on learning skill rather than what has been learned. It is on this basis that Sternberg argues convincingly that testing intelligence involves reducing the importance of knowledge and testing the ability to cope with relative novelty.

A comparison of the nature and scope of the Bio-ecological Theory with nine other theoretical approaches to understanding intelligence

Nature of assumptions	Knowledge -based	Structural Piaget/	Psycho-metric	Information processing	Genetic biological	Multiple intelligence	Modular	Contexualist	Triarchic	Bio ecological
Existence of a substantial general ability *g*	No	Yes	Mixed	Mixed	Mixed	No	No	No	Yes	No
Existence of special abilities	Yes		Yes	Mixed	Yes	Yes	Yes	Yes	Yes	Yes
Biological bases		yes	Mixed		yes	yes	Yes		yes	yes
Transdomainal processes	No	yes	Mixed	yes	yes	No			yes	No
Context sensitive		No		No	No	Yes	No		Yes	Yes
Process v product oriented	Both	Process	Product	Process	Product	Product	Process	Both	Both	Both
Role of motivation						Yes	No	Yes	Yes	Yes
Developmental	Yes	Yes	Mixed	Mixed	Mixed	Yes	No	No	No	Yes
Inductive/deductive balance	D	D	I	D	I	D	D	I	D	I
Scope of evidence										
Sociological					Mixed			Yes	Yes	Yes
Anthropological	Mixed				Mixed	Yes		Yes	Yes	Yes
Historical							Yes			Yes
Genetic			Mixed	Mixed	Yes	Yes	Yes	Mixed	Yes	Yes
Educational		Yes						Yes		Yes
Mico level processing	Mixed	Yes		Yes			Yes		Yes	Yes
Macro level processing	Yes	Yes	Yes	Mixed	Yes	Yes	Yes	Yes	Yes	Yes
Role of knowledge	Yes			Mixed				Yes	Yes	Yes
Reaction time			Yes	Yes					Yes	Yes
Positive manifold			Yes						Yes	Yes
Life course		Yes	Mixed		Mixed					Yes

Towards a Synthesis of Different Theories and Approaches

The preceding discussion has reviewed the development of the psychometric tradition and the various divergences from and reactions against it. The thesis of Spearman and the antithesis from Thurstone and the synthesis of Cattell have been reviewed, and the criticisms of the psychometric approach from Gould, Gardner and Ceci have been reviewed.

- What is the outcome of this thinking about the nature of cognition?
- What similarities are there in these different approaches?
- What is known and what is unknown about cognitive abilities?
- Is some kind of synthesis possible from this diversity?

S. Messick

Samuel Messick, the late doyen of the Educational Testing Service, wrote a detailed review of the work of Gardner and Sternberg and compared their work with the work of the psychometric tradition (Messick 1992). Messick did not have the work of Carroll to draw on in 1992, but his view of the psychometric tradition is much the same as that presented by Carroll in 1993.

Messick focused on hierarchical factor theories of ability and argued that they provide 'ground for tempering the contrasting ground of Gardner and Sternberg' (p.4). He summarised the work of Burt/Thurstone/Vernon as supporting notions that are clearly related to Gardner's logico-mathematical, linguistic, spatial and bodily-kinaesthetic intelligences. According to Messick, psychometric research provides 'support for the empirical distinctiveness and at least semi-autonomousness of six of Gardner's modules' (p.17). Like Gardner, Messick suggests that inter and intra-personal abilities may be higher order or cross-cutting processes, and he also suggests that the personal intelligences may best be thought of as a matter of cognitive style rather than intelligence. However Messick finally rejects Gardner's notion of separate abilities on the grounds that they are interrelated and interdependent rather than autonomous:

> Thus the invoking of autonomous mental modules as opposed to a hierarchical structure of abilities would appear to be counterfactual (p.17).

Messick finds in the work of Cattell a 'heirarchical structure – that is, of multiple ability complexes functioning semi-autonomously while as the same time being organised and regulated by general intellective processes' (p.9). He sees Cattell's triadic theory of intelligence as consonant with Gardner's MI theory, and notes that Gardner's thinking of logico/mathematical intelligence as a possible supra-intelligence is consistent with a second

or third level general intelligence. Messick also concludes that although Gardner rejects the notion of general abilities, hierarchical theories of ability can accommodate Gardner's theory:

> However, just because factor analytic research supports the operation of general capacities shared by or cutting across content abilities does not mean there may not be specialised resources for each content domain of the type postulated by Gardner (p.25).

Messick's Criticism of Gardner and Sternberg

Although he is anxious to reconcile Gardner's MI theory with the psychometric view of differential abilities, Messick is critical of Gardner's approach because it is criterial rather than explanatory, and reflective or rational analysis rather than empirical. In making these claims, Messick is contending for the empirical basis of the psychometric tradition and indicating a marked limitation in the resources Gardner uses in his argument, and Messick goes further by identifying a specifically political element in Gardner's argument about cognition.

One the other hand, Messick notes that Sternberg's triarchic theory emphasises general processes at multiple levels, but he argues that Sternberg glosses over 'content facets' in his model. Messick contends that the significance of the content of thinking and knowledge on the nature and quality of cognition has to be recognised. Messick describes the overall metaphor of Sternberg's theory as a business model of intelligence as mental self-management and he dismisses Sternberg's theory of intellectual styles based on political models as merely fanciful. Messick also reviews the evidence and the argument offered by Sternberg to support his triarchic theory and concludes that it is substantially 'non factual' (p.45).

Messick defends factor analysis against claims of mathematical arbitrariness on the basis of the stability of the principle of simple structure and the convergence of factor results. For Messick the approaches of Cattell, Gardner and Sternberg each have overlapping strengths and limitations, and he criticises Gardner and Sternberg for ignoring, rather than building on or adequately refuting, psychometric research:

> Hence, if Gardner and Sternberg had treated factorial theories and research on human abilities in more depth, their empirical and scholarly efforts might have systematically built upon (or undercut) these structural formulations and advanced the science of intellect in cumulative rather than an idiosyncratic fashion (p.50).

Messick calls for a model of cognition that is a comprehensive integration, and he gives particular emphasis to the integration of knowledge into a model of cognition. He focuses on the notion of intellect rather than on intelligence, and he sees intellect as a theory of both abilities and knowledge. He notes that in contrast with Gardner and Sternberg, factor theorists have related knowledge to ability:

> Factor theorists of ability have also attempted to address the intellective role of subject-matter knowledge by simply including measures of school achievement along with ability tests in factor-analytic studies, thereby directly relating subject knowledge to the ability hierarchy (p.48).

Messick also gives explicit recognition to the significance of the conative and affective aspects of performance:

> What is needed is a theoretical perspective that comprehends the interplay of abilities, knowledge, and personality in intellectual functioning (p.49).

For Messick a model of intellect is not only concerned with cognitive processes and systems. It also treats knowledge and beliefs as cognitive mental structures, and, further it treats motivations and interests as conative mental structures.

It is worth noting here that Messick is calling for the kind of 'comprehensive theory of human abilities' that Weinert said was required for a conditional or psychological account of cognitive performance. Unlike Weinert, however, Messick sees the psychometric approach as having covered a very diverse range of factors influencing cognitive performance, and he contends that ignoring psychometric evidence about intelligence and intellect is unjustified. Messick envisaged a theory and an empirical exploration of intellect that incorporated the cognitive, the conative and affective, and the social/cultural context. He also envisaged that such work could be built on a synthesis of the whole range of different approaches to cognition.

The Bell Curve Debate and Intelligence: Knowns and Unknowns

Early in the 1990s, Messick was expressing concern about the political and adversarial nature of research into cognitive abilities, and he tried to gain recognition of common ground between the positions of Sternberg, Gardner and the factor analytic approach. Nevertheless the 1990s saw even more vituperative debate about cognitive abilities in response to The Bell Curve by Herrnstein and Murray (Herrnstein and Murray 1994).

The debate involved claim and counter claim about what has (and has not) been demonstrated by scientific research into cognition. As a result of this controversy, the American Psychological Association set up a Task Force in 1994 (chaired by Ulric Niesser

and involving Ceci and Sternberg) to make clear what had been scientifically established, what was still in dispute, and what was still unknown about cognitive abilities.

The Task Force produced an authoritative state-of-the-art overview of the understanding of cognitive abilities entitled 'Intelligence: Knows and Unknowns' (Neisser, Boodoo et al. 1996). The Task Force report noted that, although a great deal is now known about cognitive abilities, the issues remain complex and in many cases are still unresolved. It noted that debate was typified by strong feelings, and that many participants made little effort to distinguish scientific issues from political ones:

> Research findings were often assessed not so much on their merits or their scientific standing as on their supposed political implications. Thus while the *g*-based factor hierarchy is the most widely accepted current view of the structure of abilities, some theorists regard it as misleading (Ceci, 1990 p.81).

After reviewing the issues, the report concluded that intelligence scores are fairly stable during development and that they predict individual differences in school achievement moderately well, correlating about 0.5 with grade point average. The Report also notes that correlations of this magnitude account for only about 25% of the overall variance, and that population levels of school achievement are not determined solely or even primarily by intelligence or any other individual difference variable:

> Successful school learning depends on many personal characteristics other than intelligence, such as persistence, interest in school, and willingness to study. The encouragement for academic achievement that is received from peers, family and teachers may also be important, together with more general cultural factors (p.82).

According to the Taskforce, IQ test scores are the best single predictor of the number of years of education that individuals complete, with a correlation of 0.55, and that in contemporary American society, the amount of schooling that adults complete is also somewhat predictive of their social status. The report asks 'Do these data imply that psychometric intelligence is a major determinant of social status or income?' and concludes that 'psychometric intelligence appears as only one of a great many factors that influence social outcomes' (p.82).

In the matter of job performance the Report notes that intelligence test scores are at least weakly related to job performance in most settings. Although IQ scores are sometimes described as the 'best available predictor' of that performance, this conclusion is qualified by the Report noting that such tests predict considerably less than half the variance of job-related measures, and that 'other individual characteristics such as interpersonal skills, aspects of personality, etc., are probably of equal or greater importance, but at this point we

do not have equally reliable instruments to measure them' (p.83). The Report notes that the predictive power of IQ tests may reflect the part such tests play in vocational entry and advancement:

> Test scores also correlate with measures of accomplishment outside of school, e.g. with adult occupational status. To some extent those correlations result directly from the tests' link with school achievement and from their roles as 'gatekeepers.' In the United States today, high test scores and grades are prerequisites for entry into many careers and professions. This is not quite the whole story, however: a significant correlation between psychometric intelligence and occupational status remains even when measures of education and family background have been statistically controlled (p.96).

After reviewing various cognitive correlates, choice reaction and inspection time measures, the Report concludes that many recent studies show that the speed with which people perform very simple perceptual and cognitive tasks is correlated with psychometric intelligence. The report also concludes that in general, people with higher intelligence test scores apprehend, scan, retrieve, and respond to stimuli more quickly than those who score lower. But this conclusion is qualified because the direction of causation is not clear:

> Do high levels of 'neural efficiency' promote the development of intelligence, or do more intelligent people just find faster ways to carry out perceptual tasks? Or both? These questions are still open (p.96).

The Report also notes that there are modest (negative) correlations between intelligence test scores and certain undesirable behaviours such as juvenile crime, and concludes that:

> In summary, intelligence test scores predict a wide range of social outcomes with varying degrees of success. Correlations are highest for school achievement, where they account for about a quarter of the variance. They are somewhat lower for job performance, and very low for negatively valued outcomes such as criminality. In general, intelligence tests measure only some of the many personal characteristics that are relevant to life in contemporary America. Those characteristics are never the only influence on outcomes, though in the case of school performance they may well be the strongest (p.83).

The Report offers this cautious conclusion overall because:

> In a field where so many issues are unresolved and so many questions unanswered, the confident tone that has characterized most of the debate on these topics is clearly out of place (p.97).

In summary then, what does this review of the notions of abilities within psychometrics and cognitive psychology show?

Conclusions from the Dialectical Process within Cognitive Psychology

This chapter has shown that psychometrics and cognitive psychology are centrally concerned with notions of ability that are distinguished from the learning of particular knowledge and skills. They are concerned with capacities that underpin particular performances and that are involved in a range of different performances. Thus psychometrics and cognitive psychology are concerned with cognitive capacities which might be loosely called generic abilities. While the nature and extent of this generality is a matter of controversy in cognitive psychology, nevertheless the notion that such capacities are general is fundamental to the field.

This review of the way the different tendencies in psychology have sought to define different abilities has shown that the different streams are not markedly in conflict and that they show a significant degree of convergence. While the attempts to theorise different kinds of ability have had significant influence in psychology and psychometrics, and they have been the basis of a broadly based movement in psychological and vocational testing, theories of differential ability have not had much influence on educational discourse and theory or on popular discourse. However, the psychometric approach to ability, and the concomitant testing instruments and programs, have become a matter of wide public concern.

Having reviewed the major lines of thinking about cognitive performance, what might we conclude from this theory and research? And how might we relate this theory and research to the various claims made about the Key Competencies?

Although there is general acceptance of the usefulness of certain conceptions of different general abilities in cognitive psychology, there is marked contention about the notion of a single general ability, and there has been particular debate about whether this general ability is innate. A good deal of this debate is appropriately described as political rather than scientific. Similarly, one might also conclude that a good deal of the debate about the Mayer Key Competencies (and the OECD key competencies) is political rather than scientific.

Although Gardner is a trenchant critic of the notion of a general intelligence, the intelligences he hypothesises are transdomainal. Gardner's multiple intelligences cannot be adduced (as Stanley seems to do) to support an argument against the notion of generic skills. Ceci makes a direct challenge to the notion of generic and transdomainal skills, or at least contends for the overriding importance of knowledge in complex performance; and he argues that knowledge and capacities cannot be disentangled. Weinert seems to support a similar position. As much as one can admire the energy and resourcefulness of Ceci's

arguments, they are not, as he acknowledges, conclusive on the nature and significance of ***g***. And Ceci's argument about the importance of knowledge does not refute the notion of transdomainal or generic capacities. While it is true that knowledge and intelligence cannot be completely disentangled, the relative importance of each in different tasks and kinds of performance does vary and can be deliberately varied to significant degrees. Cattell's notion of fluid abilities that are applied to new challenges and that are invested in crystallised abilities is important because it suggests where and how generic capacities operate. Generic capacities are those that are mobilised to learn something new, and as one becomes more competent crystallised abilities become increasingly the basis of skilled performance.

Abilities and Competencies

The fundamental argument presented here is that general abilities are significant conceptions that should be clearly distinguished from competencies. Competence is a level of achievement in a kind of task or a domain that is a result of learning, and these competencies are predominantly, even overwhelmingly, domain specific. General abilities are those capacities that are brought to bear on many tasks and that are particularly important in learning new things. General abilities are most important in assessing aptitude for learning new things, and they are comparatively unimportant is assessing competence. The burden of these claims is to dissolve the nature-of-expertise challenge to the notion of generic skills.

Ceci's does not accommodate these distinctions or refute their implications as presented here. The importance of knowledge in skilled performance is generally accepted in cognitive psychology, but knowledge-based, contextualist and the bio-ecological challenges (assiduous as they have been in their attempts at refutation) have not discredited the notion of generic capacities or abilities. The notion of a general cognitive ability has not been refuted, although the significance of ***g*** is a matter of on-going argument. The notion of a positive manifold has not been exploded, but whether or to what extent it is innate or socially determined is a matter of on-going dispute.

There are some well-attested abilities explored within the psychometric tradition, in particular the notions of verbal reasoning, abstract reasoning and spatial/visual abilities. The psychometric tradition of abilities research is largely focussed on such general abilities that can be separated from particular knowledge and skills. And this research is fundamentally concerned with aptitude to learn rather than achievement, competence or expertise.

In some respects the psychometric tradition has been inclined to exaggerate the significance of the general abilities is has focussed on, and the various challenges to this research outlined above are significant qualifications and correctives. There is no argument about the importance of knowledge in expert or competent performance, but to focus on general abilities for certain purposes is not at odds with such a view. While the importance of knowledge to expert or competent performance can be readily accepted, the notion that performances are context specific with little or no generality is not plausible and has not been sustained.

The psychometric tradition in ability research is concerned with the notion of aptitude that aims to predict the ability to learn. It is not a theory of what is to be learned, and it is not a theory of how things should be learned. Therefore the view that notions of general abilities will distract from discipline knowledge and spawn content and context-less courses in generic skills is not a necessary outcome, and seems a rather hysterical fear. One is tempted to see this fear as the response of those who cannot see the particulars they teach as a process of also developing generic skills.

We have also seen above that the psycho-biological argument for a modular structure for the mind is not strong. Brain localisation does not add up to modularity of mind, and selective absence in cognitive performance does not add up to modularity of mind. The argument for separate intelligences is not strong, and it does not seem to be getting stronger.

This review of the development of views of human abilities within cognitive psychology has shown vigorous debate about the basis of performance. In particular there has been debate about the generality or domain specificness of cognitive processes and the place of knowledge in cognitive performance, and these issues are fundamental to the debate about generic and work-related skills. Messick called for a synthesis or a reconciliation of the different lines of development in cognitive psychology to develop an integrated and comprehensive theory of cognitive performance. The analysis above has shown that such a synthesis would aim to define a small number of distinct abilities that could offer some reconciliation of general cognitive processes with specific abilities and with different domains of knowledge. Such a synthesis would be an assistance in developing notions of work-related skills, and it will be considered in the following chapter.

7

Towards a Comprehensive and Integrated Model of Performance

The Triachic Theory

The two factor and three stratum theories developed within the psychometric tradition do not explain a great deal. The bio-ecological model of embedded cognitive complexity offered by Ceci does not finally seem any more convincing than the disembedded model he seeks to discredit. It is Sternberg in his triarchic theory who has created the most useful overview of cognitive performance.

The three sub-theories of the triarchic theory clearly identify some of the major aspects of performance (Figure 8), but as Messick (Messick 1992) rightly contends about the triarchic theory, the significance of the content of thinking and the importance of knowledge are not recognised in Sternberg's model. The processes of accumulation of knowledge are part of the componential sub-theory, and the experiential sub-theory is the process of accumulating knowledge from experience that moves a task from being novel to being familiar. This movement at the centre of the experiential sub-theory, from the novel and unfamiliar to the familiar and automatized, is very important. The arrow across the middle of Sternberg's model is the process of learning, and the movement from left to right is the process of acquiring knowledge, but there are two other matters that could and should be accommodated in this model.

The Limitations of Sternberg's Model

Leaving aside the contextual sub-theory which is best thought of as the backdrop to or as encircling the other sub-theories, Sternberg's model deals only with cognition. The model does not explicitly recognise that the process of learning the unfamiliar involves both the conative and affective, and that this process is different for different kinds of task or domains. Thus the performance of the various cognitive components of Sternberg's model is shaped by conation and affect, and cognition, conation and affect interact differently with different kinds of task and domains of knowledge and skill.

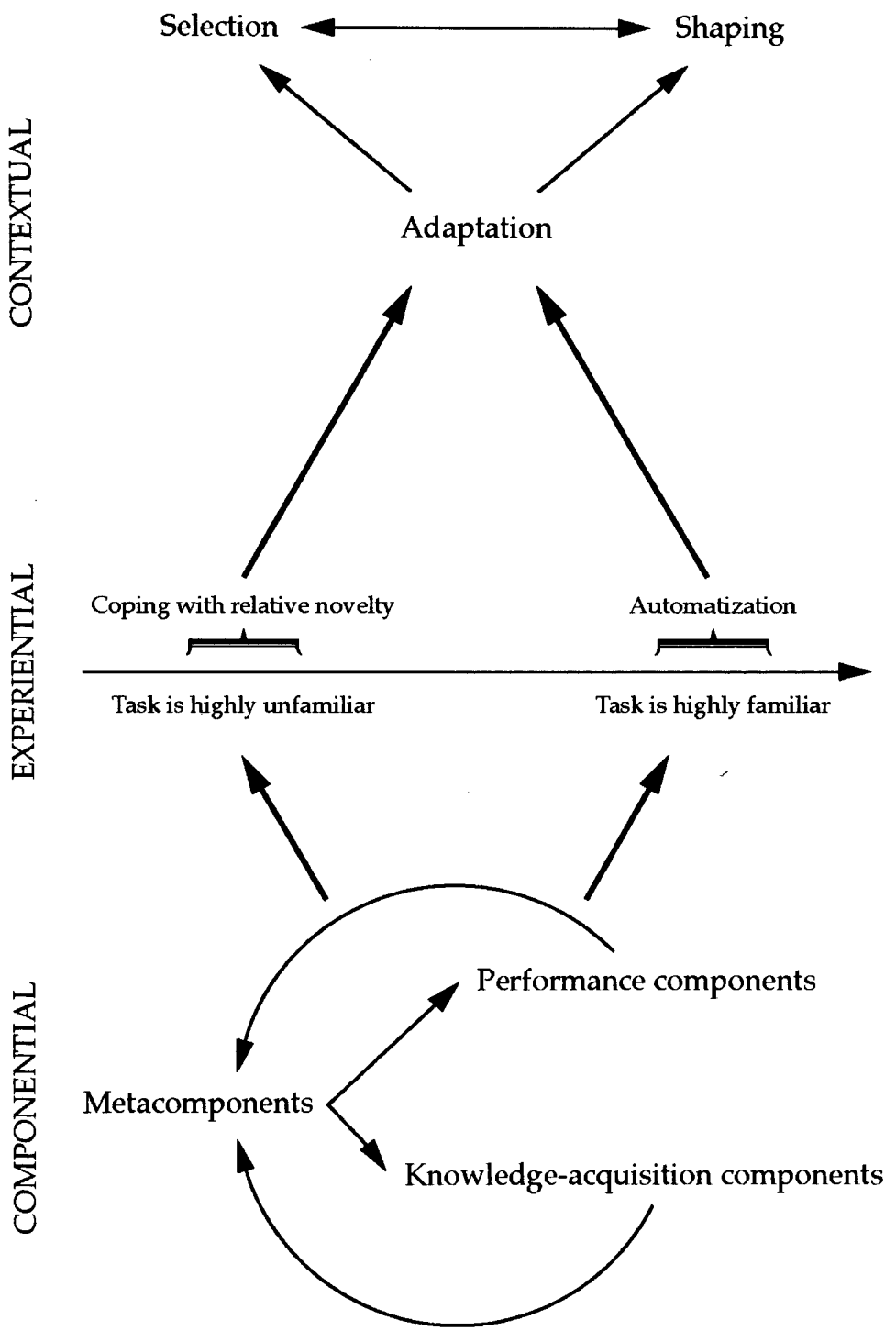


Figure 8 Sternberg's Triarchic Theory of Intelligence Revisited (Sternberg 1989)

Sternberg's model is limited in being merely cognitive. As we have seen at all sorts of points in this discussion, a theory of competence is not a more or less purely cognitive theory, and it must include attitudes and values. In the parlance of psychology, a comprehensive model of performance should recognise the interaction of cognition, conation and affect. This integration is a great challenge because it involves an integration of cognitive and social psychology, and an integration of thinking and personality.

Sternberg might contend that the interaction of conation and affect with cognition is taken for granted in his model, but there is reason to ask whether these aspects of performance can be formally integrated into his model without confusing or cluttering it excessively.

Sternberg contends that intelligence is best assessed in a situation of novelty where experience and knowledge are comparatively limited. It is reasonable to recognise that assessing abilities with tasks of comparative unfamiliarity is in itself a kind of test of learning ability, and hence a test of aptitude to learn. Thus to assess tasks at the left hand side of the experiential dimension of Sternberg's model is to test aptitude, and to test at the right hand side is to test achievement. To test aptitude is to test the efficiency of the components of Sternberg's model: that is, to test generic abilities rather than achievements. These cognitive components have a significant impact on the facility and completeness of the movement from unfamiliar to familiar: that is, they have impact on the ability to learn.

The triarchic model assumes that differences in cognitive components shape differences in facility and completeness of learning, and it assumes that differences in learning can be predicted from focussing on the operation of components doing novel or non-specialist tasks. Metacomponents are important at the left side of the experiential dimension, and they become less important as knowledge is acquired from experience. Testing metacomponents is best done within unfamiliar contexts, and metacomponents are particularly involved with the process of acquiring knowledge from experience or learning. As well as not formally recognising the importance of the feelings and wishes of the learner, the triarchic theory does not formally recognise the crucial conditioning factors of the content/domain of what is learned.

This discussion has been concerned with abilities (the componential sub-theory) which are taken to mean the capacities a learner brings to bear on a novel task or which have an impact on the performance of tasks. We have consistently encountered the notion in reviewing cognitive psychology that what a subject is being asked to learn or think about impacts on the efficiency with which cognitive components operate. What one is being asked to learn or think about is not immaterial to the operation of one's cognitive components. The particular configuration of the performance, knowledge-acquisition and metacomponents of an individual influence their performance on different kinds of task.

In this sense Sternberg's model is deficient in that it does not recognise the way in which the movement from unfamiliar to familiar interacts with and differs for different kinds of task. It is a great challenge to offer a theory of kinds of tasks, but differences in kinds of task can at least be represented in Sternberg's model.

An Adaptation of Sternberg's Model

The following adaptation of Sternberg's model incorporates, within the experiential aspect of the theory, multiple kinds of task and knowledge domains, and by incorporating a cross-cutting, vertical overlay of affect and conation, it suggests the impact of these factors on the movement from the unfamiliar to the familiar and from the novice to the expert.

This adaptation of Sternberg's model aims to make it a comprehensive and integrated model of performance, of the kind envisaged by Messick:

> What is needed is a theoretical perspective that comprehends the interplay of abilities, knowledge, and personality in intellectual functioning (p.49).

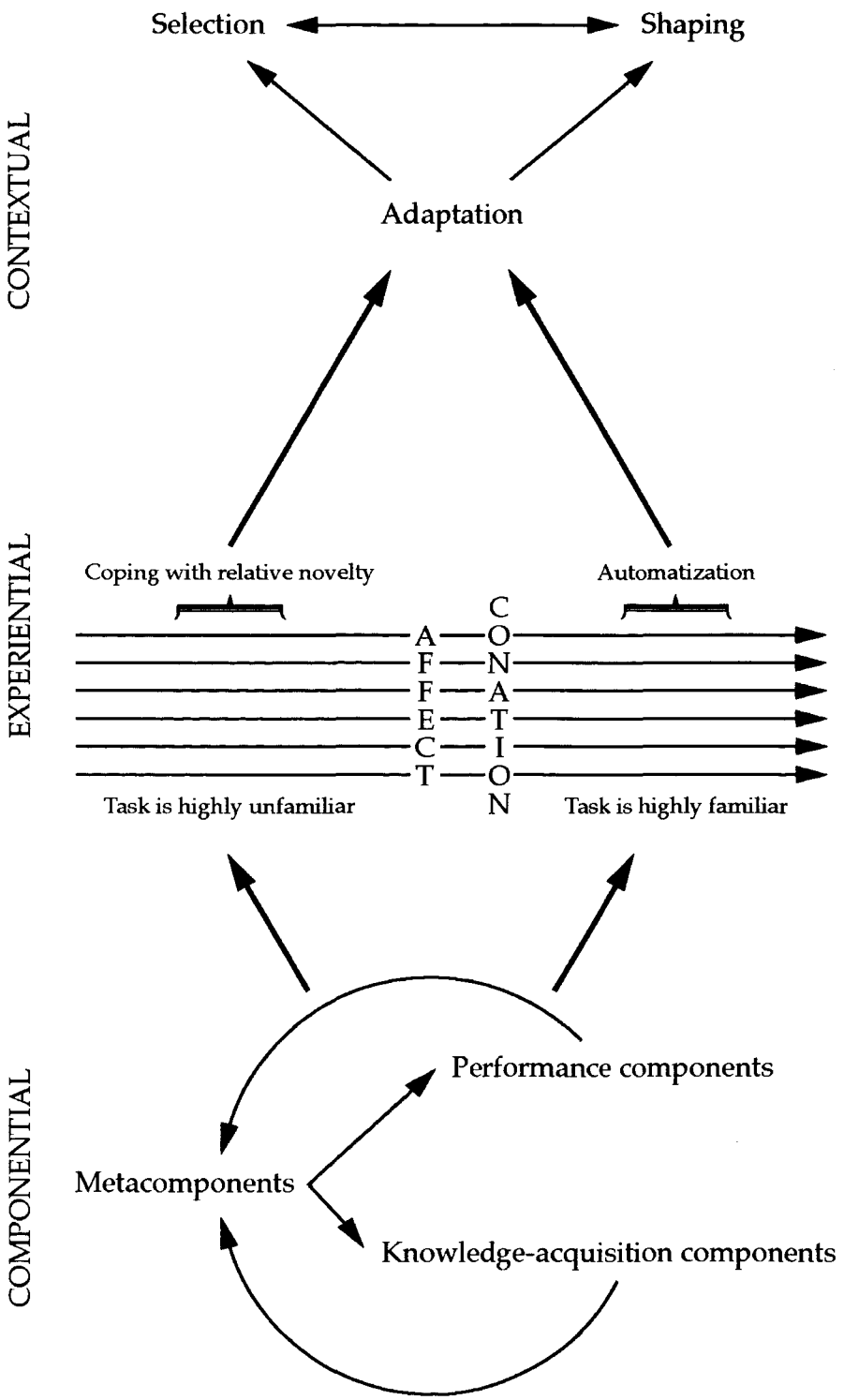
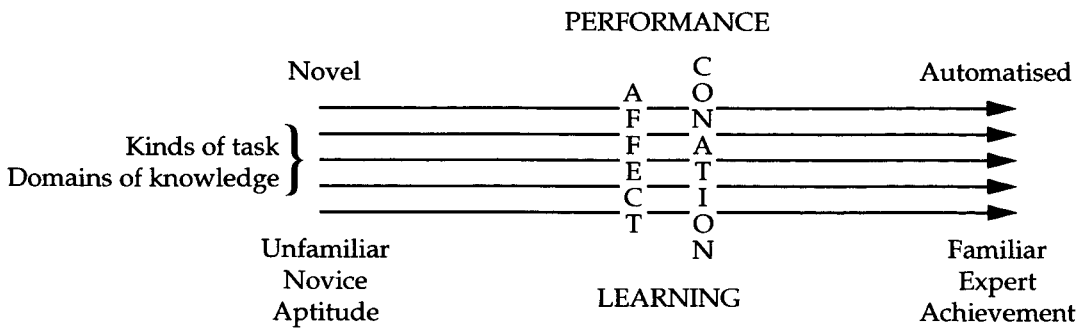


Figure 9 The Triarchic Theory of Performance Including Kinds of Task and Domains of Knowledge and the Impact of Affect and Conation

The integration envisaged in this model might be given more detailed representation in the following elaboration.

THE EXPERIENTIAL AND DEVELOPMENTAL

- moving from novice to expert and incompetent to competent
- the accumulation of knowledge and mastery of kinds of tasks



THE COMPONENTIAL AND BIOLOGICAL

- the capacities used in accumulating knowledge and the mastery of kinds of task kinds of ability

Figure 10 An Elaboration of Sternberg's Experiential and Componential Dimensions

This integration of kinds of task, knowledge, affect and conation into Sternberg's model makes it a comprehensive representation of performance.

Interest in notions of 'competence' and 'competencies' is prompted by a desire to integrate knowledge, skills, attitudes and values. Such a notion of competence involves an integration of cognition, conation and affect, or an integration of cognition with personality.

Some Notes on the Relationship of Cognition and Personality

While it might be argued that the separation of cognition, conation and affect is crude, simplistic and essentially illusory, they are actually meaningful and useful distinctions in theorising performance and in understanding work-related abilities.

As its name betokens, cognitive psychology focuses on cognition and generally excludes consideration of conation and affect. Cognitive psychology seeks to understand thought processes. It excludes consideration of conation and affect in an effort to clarify and simplify the focus on cognition, through such notions as Sternberg's cognitive components. Nevertheless there is no reason for thinking that cognitive psychologists must necessarily, or actually do, view learning and performance as impersonal information processing which is independent of conation and affect.

There have been various references to the relationship between cognition, conation and affect in this discussion, particularly in the rather loose phrases 'more cognitive' and 'more attitudinal'. It was suggested in the earlier discussion of different notions of work-related skills that different kinds of tasks and different kinds of abilities and competencies can be analysed along a spectrum of performance with the more impersonal and less attitudinal cognition at one end to the more personal and attitudinal performances at the other end.

The following diagram suggests how the more attitudinal kinds of performance have a preponderance of the affective and the conative, whereas the more cognitive is the reverse.

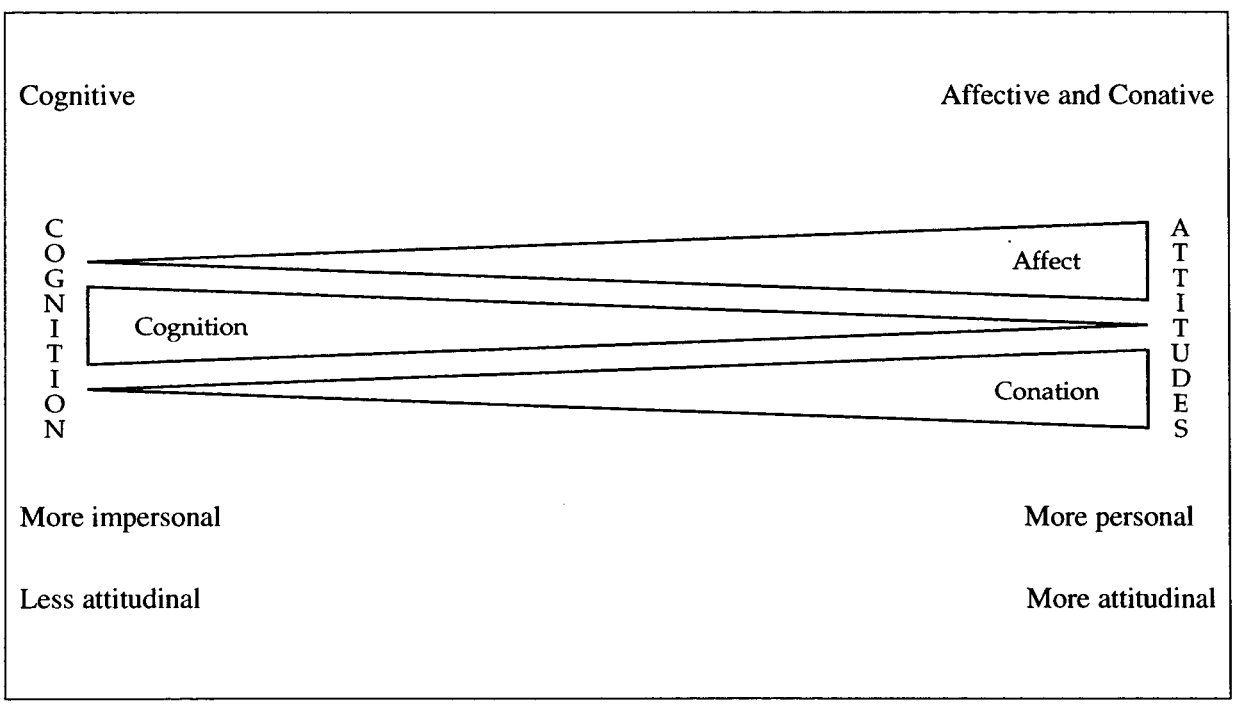


Figure 11 A Spectrum of Performance from the More Cognitive to the More Attitudinal

This spectrum of performance can be elaborated with the following distinctions or antitheses.

Table 14 The More Cognitive to Less Cognitive Spectrum of Tasks

More cognitive	↔	Less cognitive
Less attitudinal	↔	**More attitudinal**
Less personal	↔	**More personal**
Technical skills	↔	**Personal skills**
Specific knowledge	↔	**General skills**
Ability	↔	**Disposition**
Intelligence	↔	**Character**

These distinctions are related to cognitive and attitudinal spectrum of performance presented in Figure 11, and they suggest some important differences in kinds of thinking and performance.

Kinds of Thinking

These considerations lead to the notion of a spectrum of kinds of thinking and personality shown in Figure 12. This figure is an overlay or elaboration of Figure 11. It maps the relationship between different kinds of thinking and personality. In particular it attempts to interpret notions of social, personal and emotional intelligence that are being given increasing prominence in current theorising about performance.

Cognition				Personality
Formal Abstract Cognitive Processing	Social intelligence	Personal intelligence	Emotional intelligence	

Figure 12 A Spectrum of Intelligences

Among the different kinds of thinking or intelligence that have been designated by different theorists, the most important distinctions are those between notions of intelligence as cognitive processing and notions of social, personal and emotional intelligences. How might these different notions be distinguished?

Notions of Information Processing and Social Cognition

Social, personal and emotional intelligence are presumably to be distinguished from abstract/formal, impersonal and unemotional intelligence. The notion of abstract/formal, impersonal and unemotional intelligence may at first seem surprising, but there is a sense in which some traditional intelligence tests (particularly in the form of abstract reasoning tests like Raven's Matrices) are impersonal and unemotional. Abstract reasoning tests like Raven's matrices aim to be tests of formal and logical reasoning. They aim to minimise knowledge by being abstract: that is, by having no relationship to the world or to the way one usually thinks or what one usually thinks about in the world.

This abstract intelligence deal with symbols that have no relation to the human and material world, and as a result such intelligence does not deal with matters that have any emotional charge or any social or personal meaning. Abstract intelligence is unemotional and impersonal processing of information (it is what a computer can do), and hence it contrasts with thinking about socially meaningful and emotionally charged issues about which an individual may have quite personal feelings and ideological commitments.

Issues that involve social, personal and emotional intelligence have meaning in the world and they are not readily represented as formal, logical reasoning. They involve different kinds of thinking, or a range of kinds of thinking. The notion that there are different kinds of thinking is not to be seen in much of the psychometric tradition, and the greatest weakness of the psychometric view of thinking is that is has not managed to nuance the notion of reasoning into any significant differentiation.

The upshot of these considerations has been an attempt to envisage a comprehensive model of performance, of the kind represented above, that recognises different kinds of task and the way performances on different kinds of task interact with conation and affect.

Cognition, Cognitive Style and Personality Type

The spectrum of performance showing the inter-penetration of cognition and personality (Figures 11 and 12), suggests some further important distinctions. At one end of this spectrum is cognitive ability and at the other end is personality type. Because these are actually quite different categories, to represent them as different ends of a single spectrum is awkward and even perhaps seriously misleading.

We can readily envisage cognitive abilities in terms or how much of the ability an individual possesses. However, in respect to personality we more readily think in terms of kinds rather

than degrees. It makes sense to think of more or less intelligence, but it does not make sense to think of more or less personality. It is as though cognition is a continuous variable and personality is a categorical variable.

This difference might be represented by adapting the spectrum of intelligences so that personality is represented as a spectrum or a range of types with multiple lines converging on cognition. Such a relationship is represented in Figure 13. The significant feature of this figure is that it offers a middle ground between cognition and personality types that can be termed cognitive styles. These styles involve the interpenetration or overlap of different kinds of cognition, conation and affect in the manner suggested by Figure 13. This illustration shows cognitive styles as different kinds of thinking related to or reflecting different personality types. Hence, this notion of cognitive style is relating or integrating cognition and personality. If this notion of cognitive style can be integrated with notions of different kinds of tasks and kinds of knowledge then the figure represents an integrated and a comprehensive theory of performance.

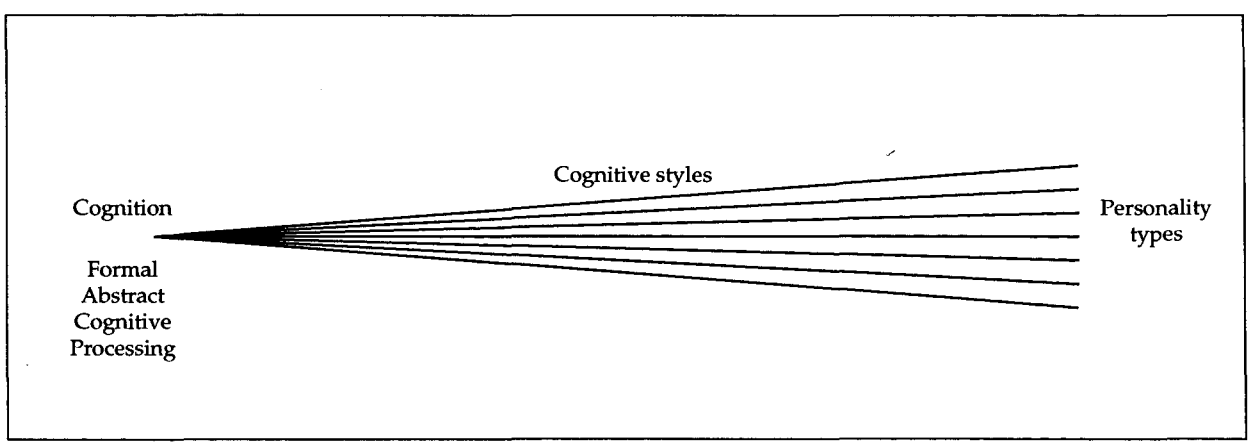


Figure 13 The Interpenetration of Cognition and Personality in Cognitive Styles

These different speculations might be brought together by returning to Weinert's paper for DeSeCo discussed in Chapter 5. With the considerations raised here in mind, Weinert's DeSeCo paper can be usefully compared with another document sponsored by the OECD, the overarching framework for conceptualising life-skills developed by the Adult Literacy and Lifeskills Project (Binkley, Sternberg et al. 2000).

DeSeCo, ALL and a Theory of Performance

Weinert's most important analytical statement in his paper is that there are three ways of defining competencies. They are as:

- cognitive processes;
- knowledge; or
- classes of tasks.

A fundamental element of Weinert's position is that knowledge is the basis of skilled performance and he sees little that is meaningful or valuable about cognitive processes that are distinguishable or separated from knowledge. This view echoes the argument commonly made over the last 25 years that it is knowledge, not generic skills, that makes the difference between novices and experts and hence that competence is fundamentally a matter of knowledge and experience.

Without dealing with the argument directly, Weinert is clearly in sympathy with this position, and it seems he has both researched and advocated it. Rather than challenging the notion of generic skills, Weinert says instead that general cognitive skills cannot be separated from knowledge. He states that there is no general model of cognitive skills, and he implies that there is, as yet, no comprehensive model of knowledge domains. As a result of these considerations, Weinert concludes that the only practical way of defining competencies is through the classes of tasks approach.

Views of General Cognitive Processes

One can readily agree with Weinert that there is no general theory of knowledge domains. Arguments about the importance of knowledge in skilled performance in recent times have been concerned to challenge views of general cognitive processes rather than developing a view of kinds or domains of knowledge. These arguments against general cognitive processes do not offer guidance for the defining of key competencies. On the other hand, one may well demur at Weinert's claim that there is no general model of cognitive skills. The Cattell model of fluid and crystallised abilities is a widely accepted general model of cognition (Cattell 1987). The Cattell model can be conflated with Carroll's three stratum model of human abilities (Carroll 1993) developed from 100 years of factor analytic research. Carroll's three strata model can be seen as a general and comprehensive theory of cognition, and one might add that Carroll can also present his work based on the factor

analysis as a class of tasks approach to defining abilities. (Carroll, as we will see, also offers a clear definition of the class of tasks approach, which Weinert does not.)

Sternberg's triarchic theory of intelligence is a more comprehensive view of cognition than that offered by Cattell or Carroll, and his model can be extended to be a fully comprehensive one by recognising the way different processes interact with different knowledge domains, and the way in which the movement from novice to expert is mediated by conation and affect.

There has been more than a century of work in cognitive psychology (from Spearman to Sternberg) devoted to understanding general cognitive processes. This work is by no means complete or definitive but it is, or should be, fundamental to the definition and selection of key competencies. This diverse body of work in cognitive psychology conceptualises and researches general cognitive abilities, and it should be, but almost always is not, the basis of defining generic, work-related abilities. This body of research needs to be refined, nuanced and related to commonsense notions of generic, work-related skills that are currently being formulated in different ways around the world.

Weinert concludes that, since a compound notion of general cognitive skills and knowledge cannot be envisaged, the class of tasks approach is the only practical basis for defining key competencies. Weinert gives almost no explanation of what he means by the notion of a class of tasks approach, nor does he claim it as a comprehensive model of performance. He implies that it is at least not theoretically unsound and he claims that it is practical. It is hard to analyse these suggestions and claims when Weinert does not explain or substantiate them.

The Class of Tasks Approach Revisited

As was argued in Chapter 4, the class of tasks approach is most readily understood as a kind of crypto-behaviourist position in which abilities are defined as classes of observable performances. These classes of tasks are taken to be observable, in contrast with elusive bodies of knowledge and unobservable cognitive processes.

Carroll formally defines an ability in introducing his compendium of factor analytic studies as classes of tasks, but he consistently suggests in his discussion of particular abilities that they are kinds of cognition as well as kinds of task. He defines abilities as classes of tasks because he is aware of the common charge made against factor analytic research that it slips into reification of statistical artefacts as cognitive processes or mechanisms.

Although the behaviourist class of tasks approach seems to offer the concreteness of observable performances, there are, as we have seen, serious challenges to be made to the

apparently concrete nature of such behaviourism. Weinert offers no explanation of the class of tasks approach he advocates and he does not subject it to the scrutiny he gives to the cognitive processes approach. In particular, he does not deal with the fundamental issue of how tasks become a class, and in what sense they are a class.

From within the psychometric tradition, however, the answers to these questions are clear. For Carroll, defining a class of tasks is a matter of the statistical properties of candidate performance that establish convergent and discriminant validity. As Weinert does not seem to take a psychometric position on classes of tasks, what he sees as the basis of a class of tasks cannot be deciphered.

Cattell's Model, Sternberg's Triarchic Theory and the ALL Framework

In contrast with the view of Weinert, it can be argued that a comprehensive model of cognition and competence can be developed from cognitive psychology. In particular, as suggested earlier, it can be argued that Sternberg's triarchic theory is a comprehensive model of cognition and can be readily extended to become a comprehensive model of performance.

As part of the Adult Literacy and Life-skills project, M. Binkley, R. J. Sternberg, S. Jones and D. Nohara have attempted a comprehensive synthesis of quite different views of cognitive ability into an overarching framework for the development of life-skills. The overarching framework (Table 15) is a matrix of cognitive processes, drawn from cognitive psychology, against commonsense notions of abilities gathered from the work-related skills movement. The process columns of this matrix incorporate the notions (outlined in the preceding chapter) of fluid and crystallised ability developed by Cattell with Sternberg's differentiation of analytical/academic, creative and practical intelligences.

The distinction between fluid and crystallised intelligences is at the heart of the Cattell model. This model attempts to account for the development of ability through the accumulation of knowledge, and to distinguish this 'knowledge' aspect of ability from generic capacities. Fluid intelligence is typified by Cattell as the kind of low verbal or non-verbal, abstract and logical reasoning tasks represented by tests like Raven's Progressive Matrices, but there is a serious weakness in this fluid/crystallised dualism, particularly as a basis for defining work-related skills.

The designation of a kind of non-verbal, abstract and formal reasoning intelligence described as fluid by Cattell is a strength of his model. The notions of spatial/visualisation and abstract reasoning are the most valuable products of psychometric research into cognitive abilities.

The weakness of the conception (in the work of Cattell and other factor analytic research) is that it effectively lumps together everything other than this non-verbal, abstract reasoning and spatial/visual processing under the heading of crystallised/verbal ability, and this crystallised ability is then explained as largely a matter of accumulated knowledge.[3] This move radically reduces the possibilities for the crucial task of nuancing or differentiating kinds of reasoning.

Table 15 The ALL Overarching Framework for Life-skills

	Crystallized Analytical (recall)	**Fluid Analytical** (reasoning, information processing)	**Practical Abilities** (standard applications- one's own life)	**Creative Abilities** (novel situations)
Communication Speaking Listening Reading Writing				
Mathematical				
Problem Solving				
Intrapersonal Motivation Metacognition				
Interpersonal Teamwork Leadership				
Using Technology				

Although the adequacy of the Cattell model as a differential ability theory can be challenged, it is more important to this discussion to recognise that the Cattell model is fundamentally at odds with the notion of a more or less comprehensive set of generic abilities. The distinction between fluid and crystallised intelligence cannot be accommodated in a view of generic abilities because fluid reasoning as defined by Cattell is too narrow a construct to adequately delineate generic abilities.

The second constituent of the ALL model is Sternberg's elaboration of his general triarchic model of intelligence into analytical/academic, creative and the practical intelligences. Sternberg has used the notion of practical intelligence in particular as the basis for

[3] The pair of Raven's Progressive Matrices is a vocabulary test of how many words a candidate can match with a synonym. This vocabulary knowledge is said to be 'crystallized intelligence'.

understanding life-skills, and he would presumably extend this thinking into notions of generic, work-related skills.

Sternberg's notion of practical intelligence aims to compensate for the absence of knowledge in his triarchic theory of intelligence by emphasising the importance of tacit knowledge in the notions of successful and practical cognition. Sternberg's triarchic theory is strong and useful, but the notion of practical cognition is not a strong or convincing extension of Sternberg's triarchic theory. The notion of practical cognition does not have the clear and definite conceptual underpinnings of the triarchic theory, and it does not result in a meaningful or useful differentiation of abilities.

One may well ask what practical intelligence is distinguished from – impractical intelligence? And one may also wonder in what sense practical intelligence is an intelligence if it is fundamentally a matter of tacit knowledge.

Leaving aside these more theoretical objections, in practical terms the ALL framework is very unwieldy. At its most parsimonious the ALL model has 24 different cells, and at its most extended it has 44. A differential ability theory that has a minimum of 24 facets is not useful because one cannot expect to envisage, validate or test 24 different dimensions.

Additionally, the ALL model does not differentiate between kinds of constructs. We are not told what similarity or difference we should see between the interpersonal and problem solving or reading, for instance. The model does not distinguish cognitive from other kinds of constructs. It seems to treat motivation and metacognition as equivalent elements of the interpersonal.

The rows of the ALL framework are drawn from notions of generic or work-related skills gathered from the various work-related constructs developed by different authorities. In one sense the rows of the ALL framework are rather commonsense, as is typical of such work-related constructs. They do not draw on and are not related to cognitive research, or any discernible theory of ability. Little rationale is offered for the selection of these dimensions in the ALL framework, and they do not amount to a coherent or comprehensive set.

The columns of the ALL framework conflate the general processing/knowledge differentiation of fluid and crystallised abilities developed by Cattell into Sternberg's broad domain differentiation of analytical/creative/practical. But there is an awkwardness and redundancy in this conflation. Fluid ability in the Cattell model is much the same as creative abilities in the Sternberg model. These dimensions could well be collapsed, but then the uneasy relationship between the different models becomes more evident, and it is likely that

in effect the Sternberg model would be collapsed into the Cattell model. It was argued earlier in this chapter that the Cattell model is not a useful basis for defining work-related skills because (like Spearman's two factor theory) it adds up to no more than the assertion that crystallised intelligence underpins any generic or context specific set of skills.

The rows of the ALL model, on the other hand, equate quite specific domains like mathematics with a broad notion of intrapersonal and interpersonal domains. The rows also include the egregiously amorphous notion of problem solving which is presumed to be a more or less general cognitive ability. This suggests that the truly generic skills of speaking, listening, reading and writing are somehow the equivalent of mathematics and technology in the ALL model, whereas it should be recognised that the macro language modes are evidently generic and that mathematics and technology are different knowledge domains.

The development of a sound model of work-related and life-skills requires a clear conceptual base. There is a significant lack of coherence and consistency in the ALL model that would limit its theoretical soundness and its practical usefulness.

With the considerations outlined in the chapter above in mind (particularly the relationship between knowledge and cognitive processes), a model of general cognitive abilities is presented below.

A Domain and Process-related Differentiation of Cognitive Ability

The previous chapter showed that for over 100 years cognitive psychologists have theorised and empirically explored notions of general underpinning abilities that are thought to be involved in all kinds of learning and performance. This research can and should inform notions of academic ability and broader notions of competencies and life-skills. In the past 30 years there has been considerable challenge to the notion of general cognitive abilities that are separate from, or separable from, the knowledge an individual accumulates about different domains. With such considerations about the relationship between knowledge and cognitive processes in mind, a model of general cognitive abilities is presented below. This model aims to accommodate both very broad knowledge domains and general cognitive processes. It also aims to offer a clear basis for understanding the cognitive aspects of competence in what is called the Domain and Process-related Model of Cognition (the DomProC model). The extension of the DomProC model of cognition into the DomProAt model of performance represents the holistic and comprehensive nature of notions of competencies.

The Domain and Process Model of Cognition

Unlike notions of competence and competencies, the DomProC sketched below (Figure 14) is a model of kinds of thinking rather than a general and comprehensive model of performance. A general model of performance recognises the significance of conation and affect as well as cognition in performance. As a model of cognition, DomProC does not seek to include conation and affect, so it excludes the conative/affective constructs such as motivation.

Nor does the DomProC include such notions as teamwork and leadership that are more attitudinal and have a significant emphasis on or admixture of conation and affect with cognition. Teamwork and leadership notions can be tested as a complex of cognition, conation and affect (that is, as a competency) but they can only be tested in comparatively complex and extended performance tests. The DomProC is concerned with the kind of thinking that underpins effective teamwork and leadership and other generic competencies, but it views these notions more broadly as particular aspects of socio-political and inter and intrapersonal reasoning.

The DomProC is a differentiated model of reasoning. It aims to synthesise content and process, or kinds of topics with kinds of thinking. It is a nested, tripartite model. At the core or centre of the model are the most generalised constructs of general cognitive ability and the truly generic skills of speaking, listening, reading and writing.

Significantly, the model interrelates cognition with language skills as the most generalised of constructs. The DomProC model views the aspects of language skills at the core as transdomainal. Viewed in this way, the assessment of language skills is also an assessment of the quality of the understanding or expression of ideas in language. [4] This relationship between language skills and general cognitive ability is a distinctive and significant aspect of the DomProC model. Rather than distinguishing crystallised verbal ability from reasoning, the DomProC recognises the sense in which thinking in language can be distinguished from domain-related knowledge and use of language. This characteristic of the DomProC rejects the notion of abstract, formal reasoning as the core of cognition. The DomProC hypothesises the core of cognition as either a complex of reasoning in language or as involving different facets of abstract and formal reasoning and informal reasoning in natural language. The DomProC puts the generic aspects of language at the centre of cognition.

[4] Other, more specialised language, skills are involved in the different domains and process of the model at the second and third layers.

The transdomainal constructs at the centre of the DomProC are given a primary differentiation at the second layer, as two very broad content domains. [5] This differentiation is based on the view that these two very broad domains typically involve different kinds of thinking as well as different topics for thought. It is based on the view that thinking about the material world and thinking about the human world involve different kinds of understanding and different kinds of thinking. This notion is extended in the third layer which breaks each of the second layer domains into four sub-components to give a comprehensive eight cell model of kinds of reasoning.

[5] These two very broad domains reflect the Finn Committee competencies of Scientific and Technological Understanding and Cultural Understanding (p.37). They also reflect the Science, Technology and the Environment (Environmental Education) and Social and Cultural Studies (Social Education) of Piper's model of curriculum (p.104).

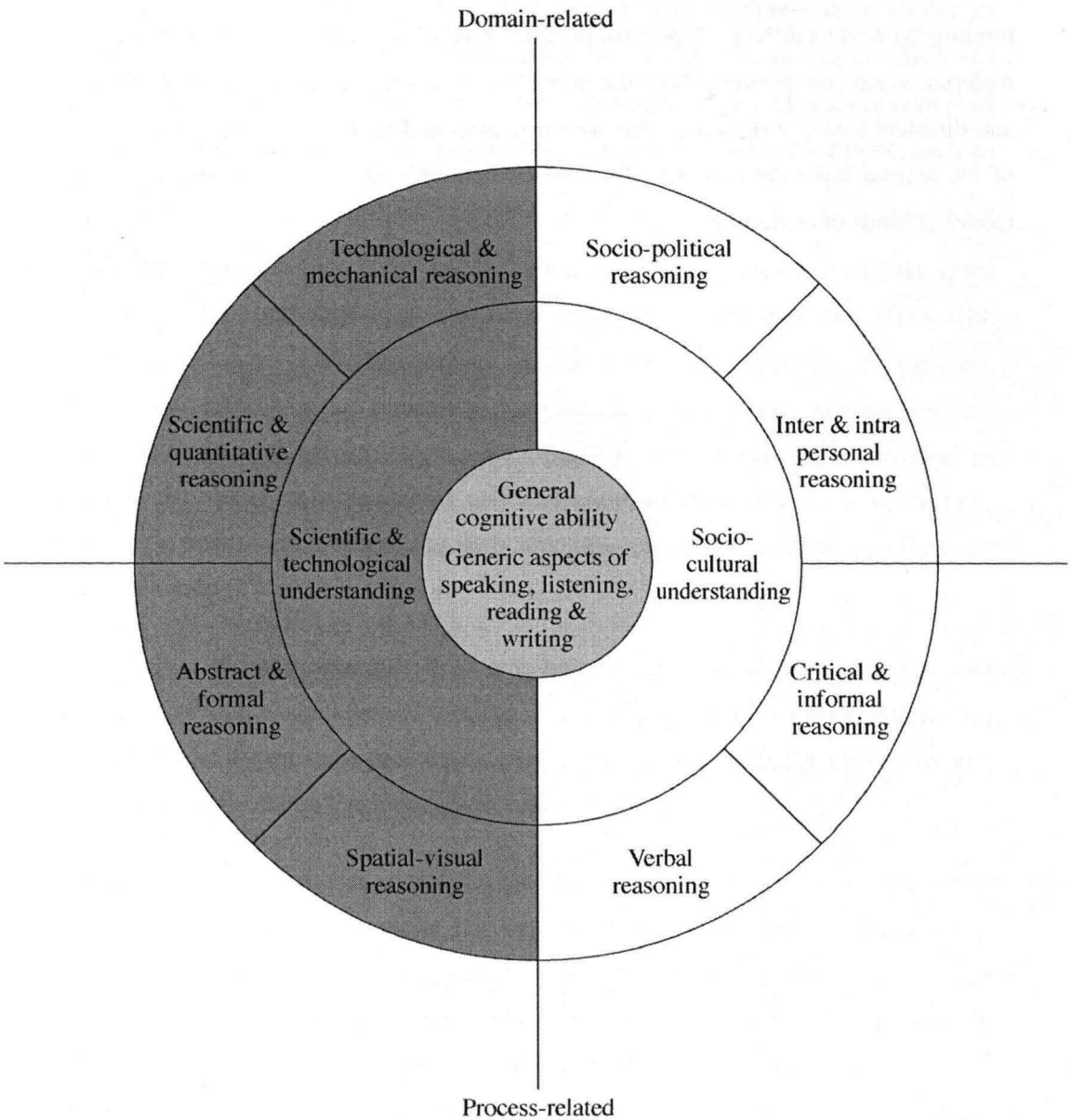


Figure 14 The Domain and Process-related Differentiation of Cognitive Ability (The DomProC Model)

Although they are all understood as kinds of reasoning, there are differences between the eight cells at this third layer of the DomProC model. The bottom half of the eight components of the third layer includes spatial/visualisation, abstract and verbal reasoning, constructs that we have seen are developed within the psychometric tradition. As such, these constructs are not domain-related. They aspire to a psycho-biological foundation, and are the most readily justified as kinds of reasoning distinct from knowledge. On the other hand, the upper half of the third layer contains scientific-quantitative, technological-mechanical, socio-political and inter and intra-personal kinds of reasoning. These constructs are viewed as kinds of reasoning conditioned by these four different content sub-domains.

The notion of critical reasoning in the bottom half of the third layer as an aspect of socio-cultural understanding may not seem to readily fit into the more process or more content related differentiation of these kinds of reasoning. But in the DomProC model, critical reasoning is understood to be particularly concerned with a kind of reasoning that is a counterpoise to abstract reasoning. Similar concerns may be raised about the absence of the seemingly generic (and ubiquitous) notion of problem solving from the DomProC model. Problem solving is not included in the model because such a transdomainal notion is more or less the same as the general cognitive ability at the core of the model. Critical reasoning can also be seen as a kind of transdomainal ability, and hence in this model it would be subsumed into the general cognitive ability at the core. If the notion of critical reasoning is seen as indistinguishable from formal reasoning and as equally applicable to the material world and to the human world, it should not appear as a separate component of this model. But critical reasoning can also be thought of as the kind of informal reasoning and value judgements that are typical of thinking about the socio-cultural world. It is as this kind of thinking, in contrast with abstract and formal reasoning, that critical reasoning is included in the DomProC.

This three layer representation of the DomProC can be seen as a four part model of reasoning shown in Figure 15.

Scientific and technological understanding Scientific and quantitative reasoning Technological and mechanical reasoning		**High verbal** Verbal reasoning Critical and informal reasoning
	×	
Low verbal Abstract and formal reasoning Spatial-visual reasoning		**Socio-cultural understanding** Socio-political reasoning Inter and intra-personal reasoning

Figure 15 A Quadripartite Representation of the DomProC Model

The DomProC Hypothesis

The DomProC model is based on the view that reasoning differs in different kinds of domains. It hypothesises that there is a significant difference in thinking about the natural and material world as compared with thinking about the human world, and that these differences can usefully be viewed as either Scientific and Technological Understanding (STU) or Socio-cultural Understanding (SCU). It is further hypothesised that these broadest of domains can each be usefully partitioned into four sub-domains or four process-related components.

The DomProC model offers to be a coherently conceptualised general model of cognitive performance. It attempts to reconcile or synthesise general processes with a domain-related view of cognitive performance. Because the DomProC is at a very high level of generality it aims to be substantially generic and transdomainal. Insofar as the DomProC model is generic and transdomainal, it is offering a model of cognition that underpins all kinds of performance, including academic, work and life experience.

The Domain, Process and Attitudinal Model of Abilities and Traits

The distinction between cognition, conation and affect is related to the distinction between abilities and traits. Table 16 is an extension of the DomProC model of cognition into a comprehensive model of performance involving cognition, conation and affect.

It can be argued that performance (as in holistic notions of competencies) always involves some mixture of cognition, conation and affect, but that the relative importance of each differs depending on the extent to which the task is more cognitive or more attitudinal (as in Figure 11). Thus, the DomProAt model is ordered from the more cognitive on the left to the more attitudinal on the right. The model distinguishes generic, domain-related and thinking skills (DomProC cognition) from practical and social skills, which are in turn distinguished from personal attributes and values.

Apart from offering a comprehensive overview of performance, the DomProAt model offers significant guidance for possible assessment, in the bottom row of the table. In this row the assessment possibilities of the three major sections of the model are analysed. The more cognitive columns are categorised as abilities that can be cost-effectively assessed in group testing.[6] The contiguous category of performance testing is understood to involve complex and holistic activities that are generally assessed as individual performances. Such assessments are generally more expensive than more purely cognitive tests, but they are commonly viewed as more comprehensive and valid. The more attitudinal category on the right of personal attitudes and values are personal traits rather than abilities, and they are not testable as abilities.

The DomProAt integrates the DomProC within a comprehensive model that allows better understanding of holistic notions of competencies and the possibilities for the assessment of abilities.

The figure of the DomProC is extended into the DomProAt in Figure 16.

[6] Cost-effective testing is presumed to be group administered, time constrained and efficiently scored. Most psychometric testing is of this kind.

Table 16 The DomProAt Model: A Comprehensive Model of Performance

More cognitive			Less cognitive & more attitudinal		More attitudinal	
Core & generic skills	**Broad Domains of thinking**	**Thinking skills**	**Practical skills**	**Social skills**	**Personal attributes**	**Personal values**
Cognitive ability Speaking, reading and writing Numeracy	Scientific and technological understanding Socio-cultural understanding	Abstract & spatial reasoning Quantitative & formal reasoning Verbal reasoning Critical & informal reasoning	Designing and planning Organising and implementing	Team work and collaboration Leadership and independence	Flexibility and adaptability Initiative and innovation	Trust worthiness Responsibility
Cost/effectively tested as an ability			**Performance assessment**		**Not testable as an ability**	

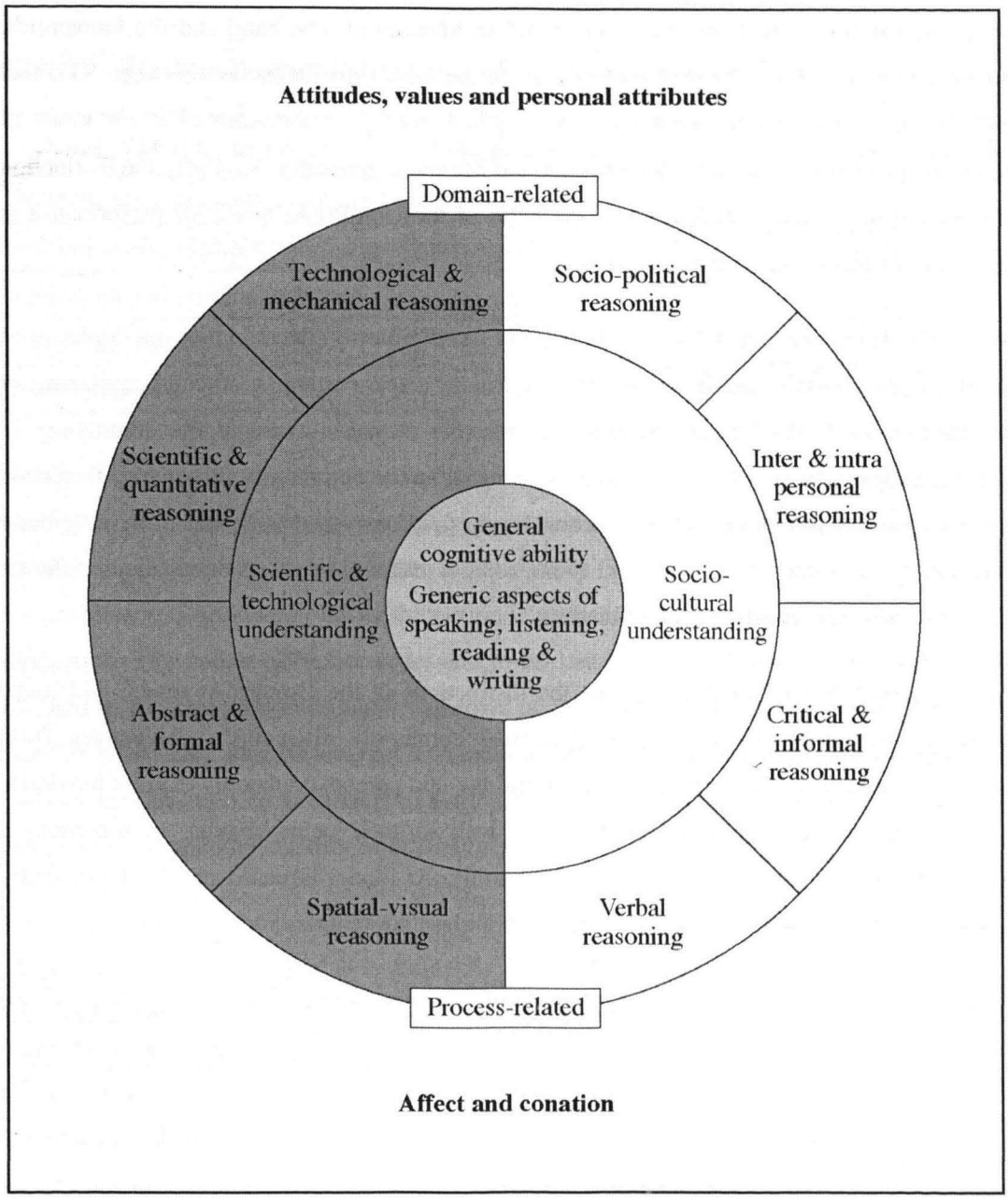


Figure 16 The Domain and Process-related Differentiation of Cognitive Ability Nested in Conation and Affect (The DomProAt model)

It would seem that the DomProC and At model of reasoning and competencies can be made intelligible to educators, business and industry and the community in general. One would expect that the DomProAt would be more intelligible to the non-specialist than Carroll's model of cognition or Sternberg's Triarchic Theory.

The dualism of Scientific and Technological Understandings and Socio-cultural Understandings will be intelligible to educators because it is related to the common distinction between mathematics, science and technology on one hand and the humanities, social sciences and the arts on the other. But in the model the distinction between STU and SCU is not a curriculum construct and it claims instead to represent different kinds of thinking about the world in both education and life more generally. Similarly the distinction between abilities and attitudes that is the basis of the DomProAt model of performance is generally understood and used by teachers.

While the DomProAt model should be readily intelligible to educators, its intelligibility to business and industry and the community in general is less certain. A crucial determinant of the usefulness of the DomProAt model is whether it makes sense to the community in general and/or whether it can be made meaningful to the community in general. A second crucial issue for the DomProAt is whether business/industry and the community in general find the layers of the DomProC and DomProAt a useful way of thinking about different abilities and tasks, as distinct from attitudes.

There is some reason for thinking that the distinctions of the DomProAt model will make sense to the community in general. Educators commonly distinguish achievement from attitude in reporting to parents about their children, and parents commonly express the desire to know about their children's attitudes as well as their achievements. In mounting a comprehensive theory of performance, the DomProAt model is based on the theoretically sound and commonsense distinction between abilities and attitudes.

However, the most important issue for the DomPro model is not whether is has commonsense appeal but whether it can be shown to have discriminant and convergent validity and whether is has predictive validity. These notions of validity are drawn from the psychometric tradition, and they involve statistical analysis of assessment data and the statistical analysis of the relationship of assessment data to actual performance in educational and work contexts. It is in this sense that the DomPro model (in the spirit of the psychometric tradition) would seek to be an empirically grounded theory.

A Model of Cognition for Particular Domains

The DomProAt is an overarching model of performance. The DomProC is a highly generalised model of cognition. The basic DomPro model can also be extended into a DomProPD for the purpose of analysis within particular domains or areas of study.

The DomProPD presented in the table below is a matrix of activity and process rows against a foundation knowledge column.

Table 17 The DomProPD Model of Cognition for Particular Domains

Kind of activity or process	Foundation knowledge of domain or area of study
Comprehension and understanding	(More convergent and objective)
Description and analysis	
Application and problem solving	
Critique and evaluation	(More divergent and opinionative)
Creative thinking	

Each kind of activity or process row in this table is presumed to take its meaning from or be meaningful in terms of the knowledge of a particular domain.[7] The activity or processes categories are not presumed to be hierarchically ordered. It is not presumed that comprehension is necessarily easier than critique or that description is necessarily easier than application. The notion of higher order and lower order kinds of activities/processes is not assumed in the DomProPD model.[8] The rows of the DomProPD matrix can be seen to be parts of an overall process, in that comprehension can be seen to precede critique and analysis can be seen to precede problem solving. This notion of precedence is not an hierarchy of complexity or order of difficulty.

The DomProPD is ordered from the receptive to the productive processes. Comprehension is convergent and creative thinking is divergent. Description is more convergent and critique is more divergent. In terms of their degree of convergence and divergence, the DomProPD processes can be understood as more objective or more opinionative.

The DomProPD is an extension of the DomProC. The different processes of the DomProPD are not presumed to apply equally to different domains or areas of study. Comprehension/understanding and Critique/evaluation are more significant activities in the humanities than in the sciences, for instance. The creative and opinionative can be readily found at all levels of activity in the humanities. Description/analysis and application/problem solving are the more significant activities in the sciences. These more impersonal and

[7] A theory of different domains is important. For practical purposes this theory may be no more than the usual academic subject areas or fields of study.

[8] An example may demonstrate the reason for this rejection of a seemingly commonsense presumption of hierarchy. If one could comprehend Wittgenstein's Philosophical Investigations it would be comparatively easy to critique it. The fundamental difficulty is in comprehending it. It may well be the case that similar comments could be made about certain scientific theories.

objective activities are more convergent, and the more divergent activities of critique and creativity may only be common at the upper reaches of the sciences.

The table below uses the objective/convergent and opinionative/divergent polarity underpinning the DomProP to separate and define the terms most commonly used in assessment activities.

Table 18 The DomProPD spectrum of Kinds of Question and Task

Kinds of question and task		
The objective		**The opinionative**
Convergent		**Divergent**
Outline Summarise Describe	Analyse Compare Explain	Discuss Comment Justify Evaluate

From the abstractions of theoretical models of cognition, the next chapter will turn to the way in which these issues were understood and played out in the real world of the ACACA response to the challenge of assessing the Key Competencies.

8

The Challenge of Assessing the Mayer Key Competencies

The ambitious assessment proposals were the most distinctive feature of the work of the Mayer Committee, and these proposals were subject to significant challenge. Whereas the theoretical challenges were outlined in Chapter 3, we now turn to look at what the Mayer proposals meant in practical terms, and how they fared in the real world of educational assessment and certification.

The history and development of pragmatic thinking about these issues will be examined through the project involving members of the Australasian Curriculum, Assessment and Certifying Authorities (ACACA), as part of the Key Competencies Pilot Phase, that was reported in *Issues and Options in the Assessment of the Key Competencies for the School Sector* (Edwards and Smith 1996).

In 1996, I was involved in two reviews about the assessment of Key Competencies that were related to this ACACA project. The first of these reviews was undertaken for Commonwealth Department of Education, Employment, Training and Youth Affairs (DEETYA) and is reported in a document entitled *Approaches to Assessing the Key Competencies* completed in June 1996 (McCurry 1996a). The second review was for the Victorian Board of Studies and resulted in a report entitled *Approaches to Assessing the Key Competencies in Victoria* that was presented in May 1996 (McCurry 1996b).

The diverse thinking about the assessment of the Key Competencies in different projects was brought together in the *Report on the Outcomes of the Key Competencies Pilot Phase* that was published in December 1996 by the Ministerial Committee on Education, Employment and Youth Affairs (MCEETYA 1996). Reviewing these documents will trace the debate about, and the responses to, the assessment proposals of the Mayer Committee.

It will be most convenient to examine these documents by first outlining the analysis I made and the conclusions I came to about the options for the assessment of the Key Competencies during 1995 and 1996. This work can then be related to the ACACA project on the assessment of the Key Competencies and the MCEETYA Taskforce report on the Key Competencies pilot

phase. All these matters will then be seen to have given rise to and shaped the trial of the whole-school assessment of the Key Competencies reported in Chapter 9.

Reviews of the Assessment Options in Key Competencies Pilot Projects

The DEETYA funded project evaluation report I developed was based on discussions at various meetings of officers from the Key Competency projects and officers of the curriculum and assessment boards in the different states and territories.

My review showed that the Key Competencies projects had raised many questions and concerns about the validity, reliability and feasibility of assessing the Key Competencies. The aim of the program evaluation was to review the various causes of concern and to see how difficulties might be obviated in developing a national Key Competencies assessment.

An Overview of the Assessment and Reporting Components of Key Competency Projects in the Various States and Territories

The review I undertook for DEETYA:

- analysed the issues and alternatives for Key Competencies assessment and reporting;
- proposed and analysed 5 models for assessing and reporting the Key Competencies (Table 20 below);
- analysed the advantages and disadvantages of the 5 models (Table 21 below);
- analysed the requirements and implications of different Key Competencies assessment and reporting purposes (Table 22 below); and
- proposed a summative, provider-based assessment as an optimal model for Key Competencies assessment and reporting (Table 19 below).

The review showed that the work which had been undertaken in the pilot phase resulted in an awkward and inconsistent set of activities in the different states and systems.

A number of other things were evident from this review. Very little of the assessment work involved the kind of direct assessment of Key Competencies by teachers, envisaged in the Mayer Report. Some systems envisaged integrating Key Competencies assessment with subject assessments. The Australian Capital Territory, Tasmania and Queensland envisaged inferring Key Competencies assessments from subject assessments or tests. There was some openly declared opposition to assessment of Key Competencies levels of performance, and there was little coherence or consistency of approach in different systems. (See Appendix 4 for details.)

National consistency of approach had certainly not resulted from the state-based Key Competencies projects, and there was considerable doubt in the various boards of studies about the validity, reliability and feasibility of the kind of provider based assessment of Key Competencies envisaged by the Mayer Committee.

The Conclusions of the Reviews

As a result of my analysis of the Mayer proposals and the responses of the various Boards of studies, I argued that the crucial issues in further developing the Key Competencies were both a sense of what is important in these constructs as well as a view about what should be given priority in a Key Competencies assessment system. I also argued that it would be quite inappropriate to develop an assessment for the purpose of changing teaching and learning, as some advocates of Key Competencies assessment argued (Lohrey 1995).

Although considerable doubts had been expressed about the validity and reliability of Key Competencies assessment from within various projects, I argued that validity and reliability are relative and contingent, and that discussion of validity and reliability has to be couched in terms of: valid and reliable for what?

I similarly claimed that it was not realistic to attempt an unstandardised provider-based assessment of individuals, as the Mayer Committee proposed, and assume that the results can be meaningfully aggregated for systems and national monitoring. I also claimed that a comparatively unconstrained and unmoderated assessment would not reach the level of hard-edged reliability that can be comfortably reported under the auspices of the state or territory accreditation agencies.[9]

On the other hand, I argued that the kind of contextualised, informal and provider-based assessment envisaged in the Mayer Report could attain a kind of soft-edged reliability and could meaningfully contribute to the transition from school to work for students as well as the process of staff selection by employers. I argued that a standardised assessment task approach was inconsistent with the Mayer recommendations and that it was neither feasible

[9] In my analysis of Key Competencies assessment I make a distinction between hard and soft-edged reliability. This distinction aims to suggest that reliability is not simply a matter of more or less reliable, and that reliability is related to the purpose of the assessment. Hard-edged reliability is needed for comparing different schools and school systems because small differences are significant in such comparisons. An assessment of an individual might be soft-edged but reliable in that it claims to be approximate rather than specific and precise. The assessment described in the next chapter is soft-edged in that a region, rather than a point on a scale, is used to describe the performance of a student. The assessment might be a single point on the scale if there is high agreement between assessing teachers, or the region might be 2 or 3 points on the scale, depending on the agreement between the assessing teachers.

nor realistic. I also argued that the Mayer committee envisaged an informal, contextualised teacher judgement as the basis of the assessment, and that such assessment could attain a soft-edged reliability without having the kind of precision that could be aggregated across providers, states or across the nation. Making these arguments amounted to a very significant curtailing of Mayer proposals.

I proposed that the practical option for Key Competencies assessment was a system where judgements are made and reported by accredited providers according to guidelines formulated and administered by ACACA Boards. The system would involve initial Key Competencies assessments by class teachers on the basis of what they observed about students in and out of class. These initial assessments from different teachers could be synthesised into an overall level score for each Competency. The scores could be reported on a certificate under the auspices of, and produced by, the provider. (See Table 19.) The assessment would not require new, formal assessment activities for students or teachers. It could take place in the course of normal class work, and an overall judgement could be synthesised from the judgements of individual teachers without complicated or time-consuming procedures. Although the proposed assessment would be reported under the auspices of the provider, the Boards could monitor the system and control the accreditation of providers. The Boards could specify the procedures to be used in making the assessments and they could review the results produced by providers.

It seemed reasonable to expect that such a system could attain the kind of soft-edged reliability that would offer meaningful information to employers about the generic skills of students in an efficient and cost-effective manner. And it seemed possible that informal teacher judgement could produce reasonably reliable information about students, and that such judgements could to be synthesised into a whole-school and cross-curricular judgement.

The sketch offered in Appendix 4 of the responses of the various education authorities in different states shows a marked contrast between the work undertaken by the various authorities and my proposals. There was actually very little work undertaken on the issue of assessment of the Key Competencies by the state projects.

Queensland, Western Australia, Tasmania and the ACT were looking for ways of adapting the subject assessments into Key Competencies assessments. South Australia proposed reducing any assessment to a dichotomous judgement. Victoria and New South Wales seemed to look most specifically at an assessment of the Key Competencies. New South Wales considered specific Key Competencies assessment tasks, but by the end of 1996 the government of New South Wales had stated its intention to not undertake state-wide Key

Competencies assessment. Victoria gave consideration to teacher judgement as a basis for Key Competencies assessment, as envisaged by the Mayer Committee, but they saw very significant problems in such a scheme. None of the systems gave any detailed thought to how the kind of whole-school or provider-based system of Key Competencies assessment envisaged by the Mayer Committee might be implemented.

I undertook a further more detailed review of these issues for the Victorian Board of Studies. This review involved more detailed analysis and elaboration of the issues dealt with in the DEETYA review.

Table 19 A Proposal by McCurry for a Summative Provider-based Assessment of Key Competencies

	An Informal, Contextualised and Provider-based Assessment of Key Competencies
Scope	• covering all Key Competencies • based on as many subject areas as are meaningful for each Key Competency
Contexts	• primarily based on subjects and other in-school activities • possible inclusion of out of school activities if they are supported by plausible evidence or can be readily attested to • possible inclusion of standardised descriptors and individual descriptions of performance
Auspices	• reporting under the auspices of the provider • accreditation by central body on the basis of specified procedures • provider quality assurance based on system guidelines • monitoring of results by a central body
Mode	• global inferences about generic skills • informal judgements contextualised within class work and other activities • generic skills distinguished from the subject or context specific skills
Method	• global impression judgements made by subject teachers • synthesis of different results into an overall level • publication of overall results to contributing teachers and review of results if necessary
Reporting	• reporting of levels required for each KC • Mayer levels used as a base, and adapted into a seven point scale on which a single point or up to three points can be reported • optional reporting of standardised and unique descriptions
Monitoring	• submission of levels to central agency • review of levels in relation to results for the system as a whole and in comparison with subject-based assessments

There are two major claims made against the notion of any generic skills or Key Competencies assessment. The first is that the notion of generic skills has little or no meaning and, more specifically, that the Key Competencies are vague and meaningless rather than generic. The second claim is that the suggestions about an assessment regime offered in the Mayer Report cannot be reliable.

The essential feature of the Key Competencies is that they are the abilities sought by employers rather than a record of achievements in the school curriculum. This implies that the Key Competencies have a significantly different emphasis and scope from most educational assessments. The Key Competencies are less cognitive, more performance-oriented and more attitudinal than most educational assessments. These differences mean that the Key Competencies cannot be readily seen as part of the curriculum, and that assessment of Key Competencies cannot be readily amalgamated with subject assessments without diluting them both.

With some temerity I claimed that the Key Competencies are not especially problematic constructs, that they are amenable to credible assessment, and that they could be assessed with the hard-edged reliability required by some purposes or the soft-edged reliability suitable for other purposes. I concluded that the type of assessment proposals outlined in the Mayer Report would not attain the kind of hard-edged reliability that could support system-level and national monitoring for public accountability or selection into further education.

I also claimed that it is unrealistic to envisage a Key Competencies assessment system as competing with or replacing subject-based grades in the school system and that it would be desirable for a Key Competencies assessment to offer different information from that offered by subject-based grades. The regimes that proposed to adapt subject-based assessments into Key Competencies assessments would either confound the subject assessments, or be limited in what they could assess. Such adaptations would entail significant changes to current curriculum and assessment practices in most states and territories. Hence, I concluded that it is doubtful whether a feasible or cost-effective system of standardised assessment tasks for assessing Key Competencies (and including all providers) could be developed. Similarly, it is doubtful whether a system of standardised assessment tasks could be developed so as to produce data that can be used for hard-edged comparisons within a system or nationally. On the other hand, the kind of contextualised, informal and provider-based assessment envisaged in the Mayer Report could attain a kind of soft-edged reliability that could meaningfully contribute to the reporting necessary for transition from school to work for students and to the process of staff selection by employers.

I also concluded that the practical option for the assessment of Key Competencies is a system where judgements are made by accredited providers according to guidelines administered by a central agency. Initial Key Competencies assessments would be made by class teachers on the basis of what they observe of students in and out of class. These initial assessments would be synthesised into an overall level score for each Competency. The scores would be reported on a certificate under the auspices of, and produced by, the provider. The level scores awarded by each provider could be reviewed and monitored by a central agency.

The distinctive characteristic of the Key Competencies is the way that they can facilitate transition from school to work for students and the selection of employees. Employers want more information in selecting employees than is offered by subject grades and there is reason for thinking that employers want Key Competency level judgements and richer information (NIEF 1995; NIEF 2000). Graded information going from summary levels and standardised descriptors through to unique descriptions and even student portfolios should be the aim of a Key Competencies assessment.

In summary, I argued that the kind of contextualised, informal and provider-based assessment envisaged in the Mayer Report could attain a kind of soft-edged reliability that could meaningfully contribute to the transition from school to work for students and to the process of staff selection by employers. I also argued that the practical option for Key Competencies assessment is a system where judgements are made and reported by accredited providers according to guidelines administered by ACACA Boards. The system would involve initial Key Competencies assessments by class teachers on the basis of what they observe in and out of class.

This proposal (summarised in Table 19) could be developed and implemented by ACACA Boards. The assessment would not require new formal assessment activities for students or teachers. The assessment could take place in the course of normal class work and an overall judgement could be synthesised without complicated or time-consuming procedures. The assessment would be reported under the auspices of the provider, but the Board would monitor the system and control the accreditation of providers. The Board would specify the procedures to be used in making the assessments and it would review the results produced by providers.

Table 20 Some Options for Summative Key Competency Assessment

Key Competency Assessment Options	Who does the assessment?	Conditions of assessment	Form of the reporting	The auspices of the reporting
1. The Mayer Proposal	• providers	• no standardisation of tasks, conditions or contexts • no moderation	• three levels	• provider auspices • centrally recorded • centrally certified? • nationally aggregated
2. The Certifying Authority Adaptation Option	• central agency provider based	• direct control of the process by central agency • possible addition of non-subject or informal assessments	• levels and/or standardised descriptors or descriptions	• a central authority
3. The Certifying Authority Standardised Option	• central agency or provider based	• standardised tasks and/or moderated procedures • possible addition of non-subject or informal assessments	• levels and/or standardised descriptors or descriptions	• a central authority
4. Provider-based Levels Option	• provider based	• levels reported • inservice for interpreting levels required • possible inclusion of descriptive material	• levels and/or standardised descriptors or descriptions	• provider auspices
5. Provider-based Descriptive Option	• provider based	• standardised descriptors and/or individual descriptions reported	• standardised descriptors or descriptions	• provider auspices

Table 21 Advantages and Disadvantages of Key Competencies Summative Assessment Options

Options	Characteristics	Possible Advantages	Possible Disadvantages
1. Mayer Proposal: Provider-based assessment centrally aggregated	• provider judgements • no standardisation • no moderation • central recording of results • aggregation of results for monitoring	• based on current assessments • not necessitating new assessments • capable of covering all Key Competencies • not requiring moderation • little increase in workload • inexpensive	• open to question on grounds of reliablility, proposed levels insufficient for provider assessment • questionable purchasing power • results cannot be meaningfully aggregated
2 Adaptation of subject assessments Option	• KC assessment based on current or adapted subject assessments	• little or no additional work for students and teachers • significant reliability • inexpensive • possibility of system and national monitoring	• little information added to current assessments • inability to deal with distinctive aspects of KCs • lower validity
3 Standardised & moderated Tasks & procedures Option	• provider judgements on standardised tasks and/or procedures • moderation of results	• higher reliability than 1, 4 & 5 • more capable of dealing with distinctive aspects of KCs than 2 • potentially cross-curricular and co-curricular • possibility of system and national monitoring	• less capable of sampling less standard skills • difficult to manage • most expensive • greater workload
4 Provider-based Levels Option	• provider judgements • levels reported • inservice for interpreting levels required • possible inclusion of descriptive material	• not expensive • capable of sampling less standard performance • potentially cross-curricular and co-curricular	• not rigorously reliable • comparatively low prestige
5 Provider- based Descriptive Option	• provider judgements • informal and observational assessment • standardised descriptors and/or individual descriptions reported	• inexpensive • capable of sampling less standard performance • richer description • possible inclusion of contexts • possible inclusion of co- and extra-curricular activities • potentially cross-curricular and co-curricular	• modest claims to reliability • comparatively low prestige • complex time consuming reporting • less definitive than 4 • differences in the quality of products

Table 22 Requirement and Implications of Different Key Competency Assessment Purposes

The table below traces the requirements and implications of different purposes and functions for a Key Competency assessment.

The implications for each of the key issues is traced for each of the three major purposes.

- The Purpose of the Key Competencies Assessment
- The Scope of the Key Competencies Assessment
- The Contexts in which Key Competencies are Assessed
- The Auspices of the Key Competencies Assessment
- The Mode of the Key Competencies Assessment
- Assessors
- Methods of Making the Assessment
- Forms of Reporting
- Feasibility and Cost/benefit

Purpose	Assessment Requirements Implications	Scope	Contexts	Auspices	Mode	Method	Reporting	Feasibility & Cost/benefit	Validity & authenticity
1 National Monitoring	requires precise measurement for system comparisons	leaning to academic/ cognitive KCs	subject-related & standardised	central agency for quality control	subject-based adaptations of current assessments or new tasks	task-based & standardised	requiring levels for aggregation	only feasible in sample studies, national adaptations not feasible	sample studies can be valid and reliable
2 Transition to Higher & Voc. Education	premium on rigorous & hard-edged reliability for equitable mass selection	leaning to academic/ cognitive KCs	subject-related & standardised	central agency for quality control	subject-based adaptations of current assessments or new tasks	task-based & standardised	requiring levels for precise comparisons of large numbers	subject-based adaptations are inexpensive & unobtrusive, but new tasks are intrusive & expensive	adaptations are not directly focused on KCs, & thus may lack authenticity
3 Transition to Employment	emphasis on selection of individuals & tolerance of soft edged reliability	leaning to less cognitive and more attitudinal KCs	more diverse contexts & more individual-ised	provider with quality assurrance procedures	inferences about the generic skills, & contextual-ised in classwork	global impressions contextual-ised in classwork	diverse or unique descriptions in order to match individuals to jobs	unobtrusive & inexpensive	directly focused on KCs and taking a wide range of evidence into account

The ACACA Key Competencies Project on Assessment and Reporting

The work outlined in this chapter so far was related to and grew out of a DEETYA funded project managed by the ACT Board of Studies entitled the ACACA Key Competencies Project on Assessment and Reporting. This project was organised around the national conference of Australasian boards of studies. The aim of the project was to promote nationally compatible approaches to the assessment and reporting of Key Competencies in the school sector, taking account of:

- different methods trialed by the States and Territories in the pilot projects;
- the interests of post school providers of education and training ; and
- the interests of business and industry.

The project brief stipulated that these objectives were to be met through the staging of at least two national policy meetings of 'stakeholders', including ACACA organisations, state and Territory education departments, independent school systems, industry and employer organisations, unions and parent and community organisations. These meetings were to consider the broad range of educational and practical issues relevant to implementing assessment of the Key Competencies.

ACACA Assessment and Reporting of Key Competencies Issues

This project began with consultation within ACACA to identify the important issues about the assessment and reporting of the Key Competencies. The list of concerns shown in Table 23 was developed as a result of these discussions. While some of these concerns were matters of policy or procedure, some are technical matters that suggest disquiet about the foundations of the Mayer proposals.

The elaborated position from the ACACA agencies stated that if MCEETYA sought assessment and reporting of student achievement in terms of Key Competencies (at the standard of reliability and validity required for certificates of senior secondary education) then it would be appropriate to ask ACACA to facilitate the development of a nationally consistent approach. The achievement of the standards of reliability and validity needed for substantial state and national comparability would require an expensive and intensive moderation process and the end use might not justify the effort involved.

Table 23 ACACA Key Competencies Curriculum, Assessment and Reporting Issues

- Responsibilities for Key Competencies, assessment and reporting
- The use of performance levels
- The integration of assessment into existing practices
- The implications of Key Competencies incorporating both knowledge and its application
- The importance of context is assessing Key Competencies
- The collection and validation of data
- The comparability of teacher judgements
- The comparability of assessments across school and VET/workplace settings
- The desired level of consistency in approach within and across States and Territories
- The purpose and value of reporting exit levels of achievement in school and non-school contexts
- The teacher workload/systems/resource implications

The Boards noted that the pilot projects had researched and developed a variety of approaches to assessment and reporting of students' achievement in terms of the Key Competencies. The Boards commented that the project work had led to suggestions that the 'meaning and significance of these purportedly generic competencies is essentially different in different contexts' (Edwards and Smith 1996).

According to the Boards, there was no clear consensus about the desirability or form of assessment and reporting in terms of the Key Competencies:

> There are instead competing, contradictory and mutually incompatible demands for assessment and reporting: some seek assessment in terms of well-defined performance levels, while at the other end of the scale others reject the notion of any form of assessment (p.7).

The Boards supported the embedding of the Key Competencies in the senior secondary curriculum. They envisaged that the assessment of achievement in terms of the Key Competencies would generally be on the basis of teacher judgments. If desired, the agencies could monitor the operation of these guidelines. According to the Boards, the reporting of a student's achievements in relation to the Key Competencies should show the contexts in which the assessments are made.

The Conclusions of the Edwards Report

The Report of Edwards and Smith on the ACACA Key Competencies Project (Edwards and Smith 1996) concluded that the stakeholders from the school and vocational education sectors, the wider community and industry have given broad support to the ACACA

minimum position as a set of educational principles and recommendations which represented a step forward in developing national cohesion in Key Competency assessment and reporting.

The report identified two major assessment and reporting methodologies that had emerged during the course of the project:

- the use of a portfolio approach to assessment and reporting; and
- the development of a standards framework (or endorsement of the existing Mayer levels) to allow reporting on student levels of achievement.

The Conclusions of the ACACA Project

The meeting of the ACACA agencies in Canberra on the 11 and 12 of March 1996 brought a phase of the Key Competencies development to a conclusion. There had been no policy decisions about Key Competencies in any of the States or Territory, and none of the ACACA boards had made decisions about the implementation of Key Competencies assessment.

The Boards had supported the general significance of the Key Competencies by agreeing they should be embedded in ACACA curriculum and that there should be an opportunity for Key Competencies assessment of all post compulsory students. Further, the ACACA agencies agreed that Key Competencies assessment should be on the basis of teacher judgement, with the Boards developing and monitoring guidelines and processes for assessment and reporting. The Boards also seemed to endorse the view that the Key Competencies should be refined for implementation. Finally, the Boards recommended that if a system of central certification were sought by MCEETYA, the Boards would facilitate a nationally consistent approach to Key Competencies assessment and reporting.

This minimum agreement suggests that a number of the issues raised earlier by the Boards had been resolved in some sense and the Boards were prepared to play a pivotal role in any future development of the Key Competencies.

New South Wales Board of Studies Papers on the Assessment and Reporting of Key Competencies

It would be fair to summarise the responses in the stakeholder meetings of the ACACA agencies to the assessment of the Key Competencies as guarded and tepid. However, it was also clear that if anything was to happen (and that was by no means clear) ACACA cast themselves as the central 'stakeholders'. (If the train left the station, the ACACA agencies were going to try to be the drivers.) It should be noted that the minimum position of the

ACACA agencies is minimal. The Boards commit themselves to no more than the development of a standards framework, if required.

An undercurrent of significant scepticism is evident in the ACACA responses to the issue of assessing the 'purportedly generic competencies'. The claim that 'the 'meaning and significance of these purportedly generic competencies is essentially different in different contexts' (p.7) leads to the further claim that:

> The reporting of student achievements in relation to the Key Competencies should show the context in which the assessments are made (p.8).

These are a pair of critical claims. The claim that 'purportedly generic competencies' are 'essentially different in different contexts' is remarkable. Both what it means, and its consequences, will be examined below, but it can be noted here that this comment about 'purportedly generic competencies' that are 'essentially different in different contexts' is tantamount to the claim that they are not generic at all. Like the relational model of Bowden and Masters, the notion that generic skills are different in different contexts makes the notion a contradiction, or a least an oxymoron. To then conclude that Key Competencies assessment and reporting should include the context seems to get very close to suggesting that the assessments will be context-specific, rather than generic.

Some of these doubts can be seen in two conceptual documents produced by the NSW Board of Studies on assessment and reporting, included as the first and third appendixes in the Edwards Report (Edwards and Smith 1996). The first is an extract from *Incorporating Key Competencies into the Curriculum* entitled: 'Nature of Key Competencies and a Key Competencies Pedagogy', and the second is a document entitled 'The Role of Context in Assessing and Reporting Key Competencies'. Both extracts are based on the earlier conceptual work of Megan Broadbent who worked on Key Competencies projects for the NSW BOS.

Under the heading 'The Nature of Key Competencies' the first appendix states that:

> Research acknowledges the reality and power of general cognitive skills such as key competencies but indicates that they always function in contextualised ways. Key competencies do not function by somehow taking the place of domain, or context specific, knowledge; nor by operating exactly the same way from one domain, or context, to another. Rather they are general tools in much the same way as the human hand is. Hands alone are not enough; we need objects to grasp. Moreover, as we reach for an object, whether a pen or a ball, we shape our hands to assure a good grip. We need to learn to -handle different objects appropriately - we don't pick up a baby in the same way as we pick up a basket of laundry. Likewise, key competencies can be thought of as general gripping devices for retrieving and wielding domain-specific knowledge, as

> hands that need pieces of knowledge to grip and wield and that need to configure to the kind of knowledge in question. In other words, domain specific adjustments will always be a part of key competencies performance (p.92).

Here we see a typical concern that generic skills will 'take the place of domain or context specific knowledge'. The metaphor of a hand for generic skills is not inappropriate. A hand is used to do all kinds of tasks. At first it will do things is a clumsy fashion and with experience it will be more dexterous. But we are then told that this 'general gripping device' is used for retrieving and wielding domain specific knowledge. This seems to imply that a generic skill is a tool for using domain specific knowledge. One wonders what happens if there is no domain specific knowledge? Nothing can be done? The metaphor of the generic hand slides to become a matter of adjusting the general device to a particular domain.

The metaphor of the generic hand is offered by Broadbent as the basis for drawing the conclusion that a generic skill is a matter of understanding a domain.

> The degree to which a key competency may be demonstrated, then, will be dependent on the degree to which the student has an understanding of the domain or context in which it is assessed (p.93).

In one sense this conclusion is self-evident, and in another it is problematic. If the conclusion is that one will be more competent the more experience one has had and the more one knows about something, then the conclusion is self-evident. However, as with the relational model of Bowden and Masters, the notion that demonstration of a key competency depends on the 'degree to which the student has an understanding of the domain or context in which it is assessed' makes the notion of a generic skill meaningless. If a generic skill can only be assessed as understanding of a domain, there can be no assessment of generic skills.

The trouble here is the notion of a competency. It is reasonable to think that the notion of competence is only meaningful within a domain and that attaining competence must entail a certain level of experience and knowledge of that domain. However, this competency isn't a general gripping device, it is a particular kind of learning and knowledge within a domain.

These ideas could have led to some questions about the relationship between the notion of competence and the generalised notions identified by the Mayer committee. The argument of Broadbent amounts to the claim that assessment is domain or context-specific, and, hence, we end up with the notion presented by the ACACA position, that Key Competencies assessments should show the contexts in which they are made.

The notion of foregrounding generic skills and even assessing them seemed to entail for some critics a process of specifically teaching them. Under the heading 'The Nature of the

Pedagogy Which is Effective in Developing Key Competencies' in the appendix, Broadbent states that:

> Key competencies can only be effectively taught within a specific context. Cognitive research has indicated that isolating general cognitive skills and teaching them without regard to their contextual configuration has proved ineffective in developing student performance. Key competencies, then, can only be effectively taught through subject-specificity. If key competencies are defined in terms of the syllabus, arise naturally from it, and receive explicit focus in the process of instruction, there is a significant probability that students will develop them (p.96).

This statement rebuts the idea of teaching of general cognitive skills themselves, but such teaching was not envisaged by the Mayer Committee. Under the heading 'Implications for Assessment and Reporting' and the sub-heading 'Assessment in Context', Broadbent states that:

> The degree to which a key competency is developed through curriculum will be dependent on the degree to which the student has accumulated knowledge in the subject. It will reflect the extent to which the student has a sense of what is relevant to the subject and understands the dimensions of the subject and the principles that underlie it.
>
> Given this, the degree to which a key competency may be demonstrated will be dependent on the degree to which the student has an understanding of the domain or context in which it is assessed (p.96).

Here again, it seems that the generic skill is the same as the mastery of subject-specific knowledge. Under the heading 'Reporting on Contexts' it is argued by Broadbent that just as Key Competencies are developed, and must be assessed, in specific contexts it is best to report them in specific contexts. It is also argued by Broadbent that:

> ... aggregation of performance levels across contexts is not a meaningful way of reporting key competency performance. This is particularly the case where there is a wide discrepancy of key competency performances across contexts. It is not meaningful, for example, to aggregate a Level 1 performance, a Level 2 performance, level 4 performance and a Level 5 performance simply as a Level 4 performance (p.97).

It does seem though, in conclusion, that Broadbent envisages some kind of subject specific assessment and reporting and some kind of aggregation:

> Reporting levels of performance in different contexts together with the aggregated level suggested by the Mayer Committee would make the process of aggregation more transparent and hence be a more informative way of reporting (p.97).

These arguments attempt to assimilate Key Competencies with 'the knowledge in a subject', as the basis of the claim that assessment and reporting of Key Competencies can only take

place within the 'knowledge in a subject'. Similar arguments are to be seen in other commentators on the Key Competencies.

The Review of the Key Competencies Pilot Program

The School Task Force of the Ministerial Committee on Education, Employment, Training and Youth Affairs (MCEETYA) appointed a Working Group on Key Competencies at the end of the pilot phase. This Working Group administered a project reviewing the nearly 80 projects funded by the commonwealth government between 1994 and 1996. The results of this work were presented to the MCEETYA Standing Committee on schools in November 1996 and published as the *Report of the Outcomes of the Key Competencies Pilot Phase* (MCEETYA 1996).

The Taskforce on Assessment and Reporting

The Task Force Report's chapter on assessment and reporting identified two broad views about the assessment of the Key Competencies:

> There are two broad views — one that the Key Competencies can only be assessed in the context of individual subjects; the other that they can be assessed in an overall, synthesised fashion (p.152).

The Report was consonant with the arguments of Broadbent in claiming the 'research literature suggests that assessment can only be in the context of discrete subjects'. Unfortunately, the Report does not disclose what research literature is referred to in this claim.

The Report recognised that such a conclusion would mean reporting within subjects, and that this would open up a problem of different assessments of what was supposed to be more or less the same thing. The Report did not canvass the notion, or draw the conclusion argued in this thesis that reporting which can only be in the context of specific subjects makes the notion of a generic skill meaningless. The Report instead considered the possibility of reporting by synthesising in broad categories of knowledge (for example, the sciences and technologies, the social sciences).

On the other hand, the Report considered the suggestion that it is desirable to aggregate teachers' judgements across all curricula and provide a global report on Key Competencies across subjects and at different levels. The Report commented that while the research literature does not appear to support such an approach, some pilot projects have pursued it and reported success:

> The work of McCurry (1996) is one example. The implications of this view are that future work needs to ask how the Key Competencies (and the levels) might be focussed and refined to make them clearer and more definite. It will then be necessary to show that it is possible for teachers to agree (more or less) in their judgments about the performance of the Key Competencies by students across the curriculum and in extra curricular activities. If this agreement can be reached then it is clear that the Key Competencies will be validated as useful and meaningful constructs.
>
> It is clear that further work is required to test this and the other approaches mentioned (p152).

I had attempted to envisage in detail what it would mean to implement the Mayer proposals, which involved developing a view about what was possible and feasible in those proposals and what was not. More particularly I had advocated the view that the Mayer Committee assumed a process of aggregation (which I described as a process of 'synthesising') of assessments into an overall view, and that a Key Competencies assessment was only meaningful and practical if it was cross-curricular and whole-school.

Contextual or Generic Assessment?

Among other things, the Taskforce Review pondered some of the important conceptual issues prompted by the Mayer proposals. According to the Report, a key issue in the assessment and reporting of the Key Competencies was the extent to which the notion of 'generic skills' is meaningful and useful:

> There is a variety of views about this matter. Some of the pilots quote research which suggests that general cognitive and affective skills (such as the Key Competencies) exist, but that they are only able to function in contextualised ways. On the other hand, in the report of a major pilot project, McCurry argues that 'all skilful performances have both specific and general elements' (McCurry, 1996) and suggests that, in this sense, assessing the Key Competencies need not present a problem of a different order from that faced by existing assessment regimes (p.152).

The Report goes on to claim some middle position between the 'purely contextual' and 'purely generic' views of the Key Competencies:

> For instance, it may be that there are differences within the set of Key Competencies as to the 'extent' to which they might be generic. Some, which are more attitudinal in character, such as Working with Others and in Teams, might be 'more' generic than are more cognitive competencies such as Solving Problems which (like critical thinking) is likely to depend on the possession of 'domain' knowledge. More research is required in order to clarify this issue (p.143).

The Problem of Assessment

Although there was ready acceptance of incorporating the Key Competencies in course designs as aims/goals or as matters of minimum competence, envisaging such generic skills as matters of instruction is more difficult, and separately assessing and reporting them seems even more difficult. Peter Hill of the Centre for Applied Educational Research offered the following pessimistic description of the situation with the Mayer Key Competencies in 1995 (Hill 1995):

> Despite prompting from business and industry and acknowledgment by schools and the training sector that the Key Competencies represent important generic outcomes of schooling, there has been little progress in achieving the recommendations of the Mayer Committee (p.1).

Hill explains this conclusion by going on to say that:

> From a curriculum perspective, ensuring appropriate attention to each of the Key Competencies has proven to be a fairly straightforward matter. They can be mapped across the various subjects taught in schools and in vocational education courses (p.1).

Hill adds that assessment difficulties have been the major impediment to development of the Key Competencies:

> Significant problems have emerged, however, when it comes to their assessment and reporting. At a general level, there is the problem of devising or adapting assessment instruments and methods to enable fair and authentic assessment of the complex, generic behaviours represented by the Key Competencies.
>
> At a more detailed level, there are problems of combining observations across different subjects or different work contexts, of establishing standards that have a common meaning across subject or work contexts, and of reconciling conflicting assessments of level of performance across subjects or work contexts.
>
> Even when these problems can be resolved, there are work-load issues arising from attempts to graft assessment of Key Competencies onto current assessment procedures. In short, serious questions regarding the validity, reliability and logistics of assessment of Key Competencies remain.
>
> These problems have led many to the conclusion that meaningful assessment and reporting of the Key Competencies cannot be realised within the context of current approaches to curriculum and assessment. Others have proposed what might be regarded as a 'Claytons' approach and have suggested that these problems can be circumvented by avoiding direct assessment altogether and simply crediting students with the Key Competencies on the basis of an 'audit' of current curricula and associated assessment procedures (p.2).

It is fair to conclude that Hill's description of the challenge of assessing and reporting Key Competencies performance has held true since 1996, and it is arguable that the impact of the

Key Competencies has been negligible as a result. Whether, and if so how, the Key Competencies performance could be assessed and reported was (and still is) the basic challenge to further development of the Key Competencies and other conceptions of generic skills in Australian education and training.

Integrated or Separate Assessment of the Key Competencies?

One response to the challenge of Key Competencies assessment was to effectively deflect the Key Competencies by 'integrating' them into courses and claiming that they are assessed within the course assessment itself. This deflection is best represented in the TAFE system response to the Key Competencies.

The Mayer committee clearly viewed the separate and graded assessment and reporting of Key Competencies performance as fundamental to their proposals. The committee presented very ambitious suggestions about the reporting of levels of Key Competencies performance, but these proposals were commonly viewed as the most disquieting and unmanageable part of the Mayer proposals by the education and training community.

In an article entitled 'Did Mayer Get It Wrong?' Megan Lilly of the Australian Competency Research Centre (Lilly 1996; Lilly 1997) offered the following principles (among others) for incorporating Key Competencies into VET sector delivery and assessment arrangements.

- Key competencies must be explicit, but not separate from technical competency.
- Key competencies must be integrated into all aspects of training.
- Key competencies should be assessed in an integrated manner.
- Key competencies should be reported in an integrated manner.

Lilly saw the Mayer notion of levels of performance as inappropriate. However, it is not clear how she envisages assessment of the Key Competencies other than that it is inseparable from, or the same as, the assessment of technical competence. Integration in this scenario means that the Key Competencies are buried in specific knowledge and skills.

While it is desirable to envisage the Key Competencies as integrated into and contextualised within a course, it is not desirable to conflate both specific and generic skills in one assessment. The result of such a conflation may be to dilute both the specific and the generic, but the most likely outcome would be to melt the generic into the specific and make the generic invisible. It is arguable that the full potential of generic skills can only be realised when they are explicitly and separately assessed and reported. But as Hill notes, the

difficulties of Key Competencies assessment are manifold, and there have been very few attempts since they were announced in 1992 to explicitly and separately assess them.

The Taskforce Review supported the view that the Key Competencies can only be understood and assessed in the context of actual subjects (and possibly, other school activities). That is, for the most part they cannot be assessed as if they were generic capacities able to be employed independently of a context.

The Report concurs with the review of Alison Wolf (Wolf 1991) based in English experience that, according to Wolf, showed attempting to assess 'core skills' during the 1980s was 'a wild goose chase':

> In a number of trials of core skills assessment over numerous years, encompassing a range of educational qualifications, it was found that there were two fundamental problems: the impossibility of assigning meaning to the core skills without a context, and the impossibility of assigning levels in a reliable fashion across subjects and activities. Even where these core skills were disaggregated and clearly defined in the limited and specific context of the youth training scheme (a pre entry level vocational course) assessors were unable to report on and assess the core skills in any consistent way across contexts. So that, for example, the need to 'decide between different courses of action' was so different in different contexts that they were unable to report with any confidence that the student had this competency. Attempts to make the competencies more concrete and specific were tried but, of course, this ultimately led to the standardisation of testing and narrowing of the constructs being assessed; the very thing that the core competencies movement wants to move away from (p.145).

The Taskforce review seemed unable to see a way forward with assessment of the Key Competencies and recommends some targeted and specific trials:

> The one thing which is absolutely clear in the difficult area of assessment and reporting of the Key Competencies is that a number of targeted studies and trials is required and that they should focus more clearly than has been possible to date on the major theoretical and practical points of contention (p.154).

However the Task Force Review does not draw the evident conclusion that either the Mayer assessment proposals were misguided or the crucial assessment matters were not satisfactorily addressed in the pilot projects.

The SBKCLAP Trial

After The Taskforce Report, DETYA invited me to offer a proposal about the school-based assessment of Key Competencies along the lines suggested by the Mayer Committee. I made such a proposal, in line with the conclusions I had presented to DEETYA and the Victorian Board, described above. The proposal was accepted by DETYA and in 1997, with my co-

worker Jennifer Bryce, I undertook a school-based assessment trial of teachers assessing the Key Competencies funded by DETYA. The proposal and the implementation of this research is the subject of the next chapter.

An Overview of Arguments For and Against Generic Skills

Various criticisms of the Mayer proposals have been traced in this and earlier chapters and it would be appropriate to now summarise and evaluate the various issues involved.

Thinking about generic or transdomainal skills should be clear about what is at issue, and what is meant by generic skills. It can be said that generic skills are by definition ones which can be deployed generally but the notion has to have a stronger foundation than a matter of mere definition. It was argued in Chapter 4 that there are three ways of exploring or defining generic skills. One way is to take the class of tasks approach and try to generalise kinds of activity, such as problem solving and critical reasoning. Another way is to explain cognitive capacities, including different cognitive components and processes, with the aim of eventually explaining the psychobiological substrate of cognition. It was argued in Chapter 4 that, however tempting it may seem to be to avoid the problems of forming a cognitive theory by defining ability as a class of tasks, such a circumscription of the notion of ability is inadequate. It was also argued in Chapter 4 that performance is an interaction of component/process/operation/capacity with a class of tasks.

It was argued in Chapters 4 and 5 that abilities are reasonably assumed to be the cognitive capacities deployed in all kinds of tasks and that we commonly think of these cognitive capacities as generic skills. These generic skills are not:

- specific knowledge and skills of a particular performance;
- what the expert knows;
- competencies, or an area in which you have to be competent; or
- a particular domain.

It is reasonable to limit the term ability to cognitive capacities, as it is to use the term domain for classes of tasks and areas of knowledge and skill. These cognitive capacities are what we mean by some notions of general skills. These general skills are those which we use in moving from a novice to an expert: they are the fluid capacities we use as novices to gain expertise. These general skills are the cognitive capacities that are deployed in all kinds of tasks. Furthermore, in assessing general skills we are particularly interested in aptitude to learn, rather than achievement or competence. According to these definitions then, generic skills are those we use to learn new things as distinct from what has been learned in the past.

Individuals differ in their capacity to do different tasks and some differences in performance hold up over many different tasks. Thus we want a theory of different kinds of task, as well a theory of different kinds of capacity or ability. Such a theory of kinds of capacity is, unfortunately, not readily to hand and, even when such a theory is developed, in psychometrics and cognitive psychology, we have seen that it is not subtle. That body of theory does not readily communicate with educators, let alone the community in general. It seems to have little impact on education, let alone on matters like the work-related skills movement.

In order to assess these general skills, one has to be careful to distinguish them from particular knowledge and skills. It is difficult to envisage what it means to assess generic skills within a particular knowledge domain and to do so would seem in danger of dissolving the generic in the specific. No scheme for the assessment of the generic within particular contexts had been developed in the Key Competencies pilot phase. To see general skills as usefully assessed for various purposes is not to conclude that it is useful to teach de-contextualised, general skills.

On the basis of these positions, developed in various ways through this discussion, it is now appropriate to review the various claims made against and for notions of generic skills. The left hand column of the Table 24 presents the negative view of generic or transdomainal skills and these arguments are then tested with rebuttal or qualification in the right hand column. Table 25 gives the affirmative response to the proposition about generic and transdomainal skills in the left hand column and these views are tested with rebuttal or qualification in the right column.

Table 24 The Negative Case against Generic Skills

Are notions of generic or transdomainal abilities meaningful and useful?	
Negative	
Proposition	**Rebuttal**
The specificity of learning and transfer of skills argument	
Learning is specific and contextualised. Evidence of transfer of skills taught in one context or domain to another is not strong.	Whether a particular skill is transferred is a question about specific skills. General skills are not particular performances which can be specifically taught and specifically transferred from one domain to another. Whether there is consistency in kinds of performance across domains is not adequately tested with experiments in the transference of specific skills.
The expert and novice argument	
Domain specific knowledge is the crucial characteristic of expert performance. Experts differ from novices in the rich and structured knowledge they possess, rather than in possessing general skills.	Recognising the importance of knowledge in all performance (particularly in competent or expert performance) does not mean that general thinking skills play no part in performance or that the notion of general skills is meaningless. Assessment of general skills is concerned with aptitude and predicting the ability to learn new things, rather with than achievements or competence.
The importance of specific knowledge argument	
General notions like problem solving or critical thinking only have meaning within specific domains of knowledge and skill. Problem solving and critical thinking are different in different domains.	Dealing with the particulars of a domain in an assessment means that the particulars of that domain are predominant rather than general skills. Some aspects of thinking remain the same across domains. The same processes and capacities are used in learning and functioning in different domains.
The inability to generalise and hence, inability to assess, argument	
Assessment generalisations cannot be made about such things as problem solving and critical thinking. Assessment of such things has to take place and is only useful within particular domains. Performance differs significantly from one domain to another.	It is not clear how an assessment of generic skills in a particular domain differs from a specific assessment of the domain. Whether useful generalisations can be made about problem solving and critical thinking can be tested.
The psychobiological argument	
The mind is structured in more or less independent modules, and hence different kinds of performance are more or less independent of each other.	Thinking involves general executive processes that operate in all kinds of thinking activity. Complex thinking involves the whole brain as a system and is not specifically localised.

Table 25 The Affirmative Case for Generic Skills

Affirmative	
Proposition	**Rebuttal**
Learning as the development of capacity argument	
Learning and education involve the development of cognitive capacities that can be used in all kinds of further learning and activity.	Learning is domain and context specific. It is particular knowledge and skills that are the basis of skilled performance.
The predictive power of generic skills argument	
We assume that education in general (and training?) can be generalised and has the power to predict other kinds of learning and performance. This ability to predict future learning is based on the capacities that are used in all kinds of performance.	Assessing specific things that have been learned in the past may predict future learning as well or better than assessing general skills.
The assessing of general skills by the minimising of specific knowledge argument	
Assessing general skills by minimising the importance of particular knowledge and skills is a process that aims to test aptitude to learn and predicts ability to learn new things.	Whether such assessment of generic skills is predictive of the ability to learn new things can be tested
The general abilities argument	
There are some generic skills – with reading, writing and speaking as prime examples, and there are some general cognitive capacities that can be usefully theorised and used. These capacities are particularly important in assessing aptitude to learn new things.	General skills like critical thinking and problem solving are unbounded and so they do not have specific or concrete meaning separated from particular domains of knowledge and skill.
The empirical argument	
Various generalised abilities (and a single general ability) have been demonstrated in studies of test performances.	The predictive power of generalised abilities and a single general ability is not strong.
The pragmatic argument	
Theories of general abilities have been developed, and they can contribute significantly to a theory of performance and the identification of key aspects of performance. Theories of general abilities are the best developed and most thoroughly researched aspects of performance.	Classes of tasks offer a concrete and specific basis for defining and assessing abilities.

The open questions
Can general abilities be made meaningful and useful? Can they be made convergent and discriminant? Do they predict future learning and performance?

It will come as no surprise to the reader that the analysis offered in this thesis claims to show that is difficult to rebut significant arguments in favour of notions of generic or transdomainal abilities.

9

The School-Based Key Competencies Levels Assessment Trial

The School-based Key Competencies Levels Assessment Project (SBKCLAP) aimed to

- test the assessibility of the Mayer Key Competencies;
- develop a circumscribed and operational definition of the Key Competencies;
- design and test an efficient means of eliciting teacher judgements of Key Competency performance;
- design and test an efficient procedure for determining an overall or whole-school assessment of the Key Competencies.

The trial component of the project was designed to explore the assumptions of the Mayer Committee, that meaningful judgements could be made about the Key Competencies performance of students by individual teachers and, that these assessments could be synthesised and reported as meaningful summaries of the Key Competencies achievements of students (McCurry and Bryce 1997). In particular, the trial aimed to see whether teachers in general education programs could make Key Competency judgements without undertaking new or different work and whether there is sufficient agreement between different teachers to make it practical to produce an overall assessment from a provider for a student.

The Trial Process

The trial involved 110 teachers in 10 schools around Australia who were asked to give a global assessment of 629 Year 11 students they were currently teaching using the Key Competencies framework developed for the purpose during the project. The teachers and the students were not expected to undertake any new or different work and the assessment was made on the basis of what teachers had learned about their students from their current work programs.

Participants in the trial were asked to develop an understanding of the Key Competencies around the middle of the year and towards the end of the year they were asked to make a decision about which of eight sub-levels they saw the trial students reaching on each Key Competency. The assessments of the different teachers in each school were synthesised into an overall assessment for a student, using a set of algorithms. (See Appendix 6.) These results

were then reviewed by a member of the school staff acting in the role designated that of the Overall Assessor.

The Key Competencies Controversy

This project set out to test certain concerns about the Key Competencies by exploring whether teachers could make judgements that are not domain specific. This issue was explored on the assumption that congruence or agreement between teachers of different subjects offers evidence about the extent to which students' performances can be usefully seen as generic and can be generalised.

At a more practical level, it was argued by critics of the Key Competencies, that teachers would not be able to make generic Key Competencies judgements, or that they would not be able to make them reliably without a new and time-consuming assessment regime. So, one part of the argument that questions the feasibility of Key Competencies assessment is based on the objection to the very notion of generic skills, whereas another part of this argument is based on the view that Key Competencies assessment is not practical or cost-effective.

This project attempted to explore these issues by designing a Key Competencies assessment regime and implementing it. It indirectly addressed the theoretical reservations about the assessment of the Key Competencies by attempting to see whether teachers find that the process of making Key Competencies judgements makes sense in practice and whether the results of Key Competencies assessments are reasonably consistent. The project directly attempted to test whether a practical and cost-effective Key Competencies assessment could be designed and implemented.

The SBKCLAP Research Procedure

The trial process involved two visits to each school by one of the researchers. Prior to the first school visit, each teacher was sent *Participant Background Notes* (Appendix 5). These outlined the areas to be covered in the first preparatory session. It was not imperative that teachers had read the notes before the session. They were provided because it was thought that some teachers might be interested to know about the project prior to the first visit.

The aims of the first school visit were to:

- give a brief overview of the development of the Mayer Key Competencies, including discussion of the climate in which they developed:
- outline propositions which underpin the approach to assessment;

- familiarise teachers with the definitions of the Key Competencies to be used in the project, including an introduction to the 'facets' used to map the Key Competencies;
- discuss the process of making global impression judgements; and
- outline the purpose and role of the overall assessor.

Assumptions Underpinning the Research

The Participant Background Notes outlined the views which underpinned the approach to assessment taken in the project. It was claimed in the background notes that the Key Competencies can be meaningfully and realistically assessed in general education programs in a minimal but useful fashion without changes to current school programs without placing substantial new burdens on teachers and students. The following propositions are presented to participants in the background notes:

1. Although in some respects the Key Competencies are unfamiliar to teachers, they can both make sense to teachers and be assessed globally by teachers.

2. The Key Competencies are part of general education; teachers know things about the Key Competencies performance of their students; and, the Key Competencies can be assessed as part of current school programs.

3. Key Competencies assessments are not in competition with, or a replacement for, subject-based assessments. They offer information that is in addition to that offered by subject-based assessments.

4. Key Competencies assessments are aimed particularly at offering information about students who are in transition to work and they touch on skills and attitudes not always or readily included in subject assessments.

5. The most realistic way of reporting summary judgements of Key Competencies performance in general education is on the basis of global judgements made by individual teachers about the general skills and attitudes of the students they teach, with these individual judgements being subsequently synthesised into, and reported as, overall school levels.

6. Judgements about Key Competencies performance should be broad and general inferences about the way students are likely to perform in the future, rather than judgements about students' performances in specific courses. Because they are general or generic, Key Competencies judgements can be synthesised into collective school judgements.

These positions were presented to participating teachers as a way of placing the activity in a context.

Adaptation of the Mayer Key Competencies

While the Mayer Committee conception of the Key Competencies was the basis of this project, the Key Competencies were re-ordered according to whether their conception was seen as more cognitive rather than attitudinal (in the spirit of Table 15). In addition they were more precisely defined for the purposes of assessment by means of a small set of brief descriptors.

Collecting, analysing and organising ideas and information was interpreted as research that has a particular emphasis on analytical skills. Communicating ideas and information was located specifically in the language skills of speaking and writing. There seemed no way to make Using Mathematical ideas and techniques generic and it was represented as understanding and using mathematical knowledge.

The notion of assessing attitudes was rejected by the Mayer committee but, in envisaging how competencies such as Planning and organising activities or Working with others and in teams are to be assessed, it makes sense to see them in terms of kinds of approaches or attitudes rather than trying to assess then as particular performances. These more attitudinal constructs are best thought of as matters of cognitive style or matters where cognition and personality overlap (in the spirit of Figures 11 and 12).

The solving problems Key Competency had been widely criticised for its vague generality but if it were seen as an approach or attitude to solving problems, it was thought to make its assessment more feasible in the trial. Using technology poses much the same problems as Using mathematics. It is reasonable to ask what technology is at issue and to foresee difficulties of access and equity. On the other hand, Using Cultural Understandings, the most controversial of the Key Competencies had been so partly because some had seen it as dealing with socio-political value judgements. (NLLIA 1994; Rumsey and Hannan 1996) To circumvent this problem, it was interpreted in a more cognitive way in this project. These issues are discussed at greater length in the background notes (Appendix 5).

The principled reasoning for explicitly introducing attitudes into the Key Competencies assessment was to recognise that competence in employment is usually defined as involving knowledge, skills and attitudes. The briefing notes stress that the kind of assessment of

attitudes envisaged for this project is an assessment of personal characteristics, rather than an assessment of social conformity or ideology.

The resulting re-ordering of the Key Competencies is given in Table 26, and this rendition of the Key Competencies is that referred to in the trial outlined below.

The background notes pointed out that the first four Key Competencies (according to the conception developed by the project) emphasise kinds of thinking skills - the same sorts of thinking skills that are the basis of the school curriculum, or, more precisely, the skills assessed in school grades. The second four can also be thought of as thinking skills, but the emphasis in the descriptions has been moved away from knowledge and thinking skills to give more emphasis to attitudes and personal qualities. The second four Key Competencies (as interpreted in this trial) were asking for assessments that take into account not only the capacity of students to think about things and socio-cultural situations but also the way they feel about, approach and interact with things and socio-cultural situations.

The Facet Descriptors

In order to achieve a coherent notion of each Key Competency so that teachers would be assessing the same aspects of a student's performance, it was necessary to define the Key Competencies more specifically. This was done by developing what were called facet descriptors which aimed to identify aspects of the Key Competencies which were each different, yet a part of the coherent whole. The facets of each Key Competency thus aimed to:

- map the Key Competency and indicate the issues to be taken into account in assessing it; and
- give a means of reporting the aspect of performance that is typical of the student.

The facets were intended to be:

- parts of a general but coherent notion;
- a different aspect from the others in the group for that Key Competency; and
- a key distinguishing feature of performance on the Key Competency.

Table 26 The Key Competencies as Interpreted for SBKCLAP.

1 Collecting, Analysing and Organising Information

The capacity to locate information, sift and sort information in order to select what is required and present it in a useful way, and evaluate both the information itself and the sources and methods used to obtain it.

2 Communicating Ideas and Information in Speech and Writing

The capacity to communicate effectively with others using speech and writing.

3 Using Mathematical Ideas and Techniques

The capacity to use mathematical ideas, such as number and space, and techniques, such as estimation and approximation, for practical purposes.

4 Using Cultural Understandings

The capacity to understand and interpret socio-cultural ideas, and to reason plausibly about personal and inter-personal issues.

5 Solving Problems

The capacity and the attitude needed to apply problem-solving strategies in purposeful ways, both in situations where the problem and the desired solution are clearly evident and in situations requiring critical thinking and a creative approach to achieve an outcome.

6 Using Technology

The capacity and the attitude needed to apply technology, combining the physical and sensory skills needed to operate equipment with the understanding of scientific and technological principles needed to explore and adapt systems.

7 Planning and Organising Activities

The capacity and the attitude needed to plan and organise one's own work activities, including making good use of time and resources, sorting out priorities and monitoring one's own performance.

8 Working with Others and in Teams

The capacity and the attitude needed to interact effectively with other people both on a one-to-one basis and in groups, including understanding and responding to the needs of a client and working effectively as a member of a team to achieve a shared goal.

Table 27 The First Three Key Competencies as Interpreted for SBKCLAP

FACET DESCRIPTORS & STANDARDS SCHEMA FOR THE SCHOOL-BASED ASSESSMENT OF KEY COMPETENCIES LEVELS IN YEARS 11 & 12

Taking into account the Facets for each KC, make an on-balance judgement about the level of achievement usually attained by a student.

	Key Competency Facets ***Aspects of the KC & descriptors for reporting*** *Which facet is the typical & distinctive strength of a student?*	**Level 1** ***Basic Achievement in Grades 11&12*** (Codes B and B/M) At this level students can usually	**Level 2** ***Medium Achievement*** (Codes M and M/H) At this level students can usually	**Level 3** ***High Achievement in Grades 11 & 12*** (Code: H and H+) At this level students can usually
1.Collecting, analysing & organising ideas & information	*A comprehending and interpreting ideas and information* *B analysing and evaluating ideas and information* *C synthesising and developing ideas and information*	•deal with basic and routine research tasks •comprehend basic information •organise basic information •assess basic information as directed	•deal with some more complicated research tasks •comprehend, & analyse basic info. •organise some more complicated information. •assess basic info. independently	•deal with complicated research tasks •comprehend and analyse more complicated information •organise more complicated info. •assess more complicated information
2a. Oral Communic-ation of Ideas & Information *** **2b. Written Communic-ation of idea & information**	*aA speaking in less formal situations* *aB speaking in more formal situations* *aC using speech to interact with other people* *** *bA writing clearly and accurately* *bB writing confidently and fluently* *bC using writing to explore ideas and information*	•speak precisely and appropriately on some topics •appreciate the demands of speaking to some purposes and audiences *** •write on some topics clearly and coherently •appreciate the demands of writing for some purposes and audiences	•speak precisely and appropriately in a range of ways on some topics •appreciate the demands of speaking for a range of purposes and audiences *** •write complicated ideas clearly and coherently •appreciate the demands of writing for a range of purposes & audiences	•speak precisely and appropriately in a wide range of ways on a range of topics •appreciate the demands of speaking to a wide range of purposes and audiences *** •write more complicated ideas clearly & coherently •appreciate the demands of writing for a wide range of purposes and audiences
3. Using Mathemat-ical Ideas & Techniques	*A using mathematical techniques to represent and analyse ideas and information* *B understanding mathematical principles and procedures* *C applying mathematical knowledge to problems*	•use basic maths methods to analyse data •use basic maths reliably and efficiently •appropriately select and apply basic maths techniques	•use some more complicated maths methods to analyse data •use some more complicated maths reliably and efficiently •appropriately select and apply more complicated maths techniques	•use complicated maths methods to analyse data •use complicated maths reliably and efficiently •design, interpret and evaluate maths activities

Table 28 Key Competencies Four to Eight as Interpreted for SBKCLAP

FACET DESCRIPTORS & STANDARDS SCHEMA FOR THE SCHOOL-BASED ASSESSMENT OF KEY COMPETENCIES LEVELS IN YEARS 11 & 12

Taking into account the Facets for each KC, make an on-balance judgement about the level of achievement usually attained by a student.

	Key Competency Facets *Aspects of the KC & descriptors for reporting*		**Level 1** *Basic Achievement in Grades 11&12* (Codes B and B/M)	**Level 2** *Medium Achievement* (Codes M and M/H)	**Level 3** *High Achievement in Grades 11 & 12* (Code: H and H+)
4. Using Cultural Under-standings	*A comprehending and interpreting social issues* *B developing a coherent position on social issues* *C using cultural understanding to achieve goals*		•comprehend some social issues •take a reasoned position on some social issues •make some use of cultural understandings	•comprehend some more complicated social issues •take a well-reasoned position on some social issues •make use of cultural understandings	•comprehend complicated social issues •take a well-reasoned position on a range of social issues •make effective use of cultural understandings
5. Solving Problems	*A shows focus and persistence* *B shows independence and responsibility* *C shows initiative and creativity*		•show focus, independence, &initiative in responding to basic problems and challenges	•show focus, independence, & initiative in responding to some more complicated problems and challenges	•anticipate & show independence & initiative in dealing with complicated problems & challenges
6. Using Technology	*A .seeks technological experiences and challenges* *B .is thorough and persistent in dealing with technology* *C .can understand technological systems*		•approach some technology positively •deal with some basic technology	• approach a range of technology positively •deal with some more complicated technology	•approach a range of technology with energy and application •deal with more complicated technology
7.Planning & Organising Activities	*A identifies goals, priorities and strategies* *B implementing plans and strategies* *C seeks planning and organisational challenges and experiences*		•plan & organise some basic and defined activities	•plan & organise some more complicated and defined activities	•initiate, plan & organise complicated activities
8. Working with Others & in Teams	*A adapts to group expectations* *B leads or facilitates group processes* *C assists and supports others*		•take a role in a team and can support other team members	• contribute well to team work and can usefully support other team members	• contribute significantly to team work and can effectively promote other team members

The facets of each Key Competency are outlined in Tables 27 and 28. The criteria for defining a Key Competence are related to and reflect conventional psychometric and factor analytic attempts to define constructs, outlined in Chapter 6.

Some other projects had sought to define the Key Competencies but these definitions were usually general process sequences (Ryan 1997). The approach taken in the facet descriptors for SBKCLAP aimed at scoping statements that sought a balance of breadth and specificity. Other projects that gave any attention to assessment issues focussed most attention on the issue of defining levels of performance, rather than the constructs themselves.

The SBKCLAP did offer three briefly described levels of performance but the process of defining standards was not treated as a significant issue. The SBKCLAP approach might be described as the within-school comparative standards model, which might be contrasted with an alternative that bases itself on the definition and application of absolute standards across a system. These two models are compared in the Table 29 below. According to this analysis, the major weakness of the within-school comparative model is the way that it requires some students to be stigmatised as weak, because it sets out to identify those students who stand out either because they are strong or weak. This flaw might be ameliorated by making the comparative model only comparatively positive rather than positive and negative. In this situation, none would be stigmatised as weak, all would be assumed to be at level 1, and the assessment would be about positive strengths in students. Such an interpretation of comparative assessment would limit discrimination among the students within a school. Discrimination might be improved by both expanding the score range and including the mean and standard deviation for all students assessed on the report (McCurry and Bryce 1997).

Table 29 A Comparison of the Absolute and the With-in School Comparative Models of Assessment

	Absolute standards model	Within-school comparative model
Approach to standards	emphasises described levels of performance aspiring to common standards across the system	comparative or normative standards taking definition from teachers expectation within a school
Advantages	aims at between school comparability does not require that some must be stigmatised as comparatively weak	does not entail the production and mastery of a described scale need entail little development costs; in-service support; or on-going standards monitoring and moderation is inexpensive and low maintenance
Disadvantages	would entail development of scales, in-service support; and on-going standards monitoring and moderation; would be comparatively expensive and high maintenance	entails some students being stigmatised as comparatively weak; does not offer comparability between schools and across the system; may not be as highly valued as a comparable assessment
Approach to reliability	may be more reliable because of defined standards	may be more reliable because of tacit consensus among teachers
Impact on students	need not require some to be comparatively weak	requires some to be comparatively weak

242

The school-based comparative assessment model was chosen for SBKCLAP because it is the only inexpensive and undisruptive option for KC assessment. The trial was to determine whether this within-school standards method has the potential to be valid, reliable and cost-effective.

Assessment Using Global, Impression Judgements

Teachers were asked in the trial to make judgements about students that were not based on specific tasks or instances of behaviour. Rather, they were asked to make judgements that were overall and synthesised impressions which took into account as much evidence as the teachers could register. The generic nature of the Key Competencies was stressed in the background notes to participants. It was pointed out that teachers would gain most of their knowledge about students from subject classes but that the judgements to be made in the trial were concerned with what a teacher takes to be general about the abilities or performance of a student. That is, the judgements were to be inferences about what is typical of the student in most subject areas and other activities. The judgements for this project concerned students in Years 11 and 12 general education programs. Teachers were told to base their judgements on their knowledge and experience of students at the appropriate level generally and on what they know and expect of students at that stage.

Using the Assessment Framework

The Mayer report proposed that Key Competencies be assessed over three levels of achievement. These levels were used in the standards schema for this project: Level 1 (basic achievement at grades 11 and 12), Level 2 (medium achievement in grades 11 and 12) and Level 3 (high achievement in grades 11 and 12). To facilitate the assessment, level descriptors were written and the levels were broken into 'level codes' which enabled teachers to assess using an eight-point scale. The level codes are outlined below.

Teachers were encouraged to make their judgements by thinking of how some students stand out, either because they exceed or do not meet usual and reasonable expectations. Students who do not meet basic expectations are judged to be below level one whereas those who clearly distinguish themselves as high level performers are at level three. So most students will fall in levels one and two on a particular competency because they meet basic expectations but are not clearly distinguished.

Table 30 The KC Assessment Level Codes

Not Yet at Level One	NY
Not Yet/Basic	NY/B
Level One *Basic Achievement*	B
Basic/Medium	B/M
Level Two *Medium Achievement*	M
Medium/High	M/H
Level Three *High Achievement*	H
High Plus	H+

The process was further outlined in the following terms:

1 If a teacher is tempted to think that a student is particularly strong or particularly weak in a Key Competency, they are faced with a decision between Levels 3 and 2 for the stronger student, and Not Yet Level 1 and Level 1 for the weaker student.

2 If a student is not clearly distinguished then the teacher is deciding between Levels 1 and 2.

3 In deciding between Levels 1 and 2, teachers are looking for signs of strength that raise candidates above the basic expectation.

In briefing participants, significant attention was given to the facet descriptors but the levels of performance were given little attention. The guidelines offered to teachers for making the assessments were very concrete. They were based on the notion that teachers have expectations about what is possible and what can be expected of students at a certain level and that these expectations are quite generally shared. The notion of absolute standards manifest in described levels of performance was eschewed by the project because such attempts at defining absolute standards can be inconsistent and unconvincing. They place the proficiency scale at the centre of the process, and they give teachers the problem of making sense of the scale and applying it. The aim of this trial was to get as much agreement between participating teachers as possible with as little stress as possible. Rather than having

participants worry about the meaning of proficiency scales, the project attempted to make participants feel confident that they knew what was at issue and could make such judgements.

The Role of Overall Assessor

On the first school visit, the role of overall assessor was outlined. In most cases an appropriate staff member had already been selected - usually the Year 11 co-ordinator, who was familiar with staff and students at that level.

Initially it was planned that the overall assessor would review teachers' judgements with the aid of a computer program. For various reasons it was later decided to follow the same process using pen and paper and to leave the implementation of the software until after the completion of the assessment. The computer application for the whole-school assessment and reporting of generic skills was subsequently developed in the NIEF Portfolio project (NIEF 2000) and the VCAA assessment of Key Competencies project (McCurry and Bryce 2002).

The main aim of the overall assessor in this scheme was to bring together the judgements of teachers into an assessment of a student that would range over no more than three points of the eight-point scale. This was to be done, not by averaging, but rather by taking into account the perspectives of different teachers and noting comments or other indications they had made. Sometimes discrepant assessments would need to be reconciled by discussion with teachers and it was recognised that sometimes reconciliation might be impossible, in which case no assessment would be given for that student on the particular Key Competency.

By the end of the first school visit it was expected that teachers would have gained a clearer conception of the Key Competencies and familiarity with the assessment scheme that they would use at the end of the year.

Selection of Students to be Assessed

It was expected that in a school with 80 - 100 Year 11 students it would be reasonably easy to find a group of 6 or 7 teachers who taught or knew through other activities most of these students. This expectation was naive. Each student was to be assessed on the Key Competencies by at least four different teachers. In reality, setting this up was extremely complex because there are so many different units at Year 11 level taught by different teachers, so there were very few teachers who had taught or 'knew' a large body of the Year 11 students.

Because of this complexity some changes had to be made and some teachers had to be recruited to participate in the assessments even though they had not attended the initial briefing session. (Unintentionally, this provided an opportunity to test the clarity of the materials - even those teachers called in 'at the last minute' were able to come to terms with the procedure and undertake the assessments.)

In the end, some students were included who had been assessed by three rather than four teachers.

The Second School Visit

The conception of the overall framework did not change from the time it was presented in the first school visit. However, some modifications were made to the facet descriptors in the light of teachers' suggestions. There were three facets for each Key Competency and it was decided to ask teachers to assess each facet (as well as indicate a student's strongest facet). This was for research purposes - to see, for example, whether one facet of a Key Competency tended to get higher gradings than the other two.[10] It was not intended that this be a part of the final assessment model.

The second school visit involved meeting with teachers and ensuring that they understood the assessment framework. In some schools this also involved meeting with the overall assessor and confirming which students were to be assessed.

Because the process for selecting schools took more time that expected, there were three schools where the first and second visit had to be collapsed into one. This meant that teachers did not have much time to think about their students in relation to Key Competencies, but there were no obvious difficulties at these schools.

The Third School Visit

Teachers were given a date in November by which time all assessments were to be completed. The overall assessor supervised the return of all completed assessment documents to ACER by that date. There was a short amount of time in which to process the assessment materials into a format that could be checked by the overall assessor - a sheet for each student which showed all assessments given by teachers. Adjustments to the assessments were not made at this point.

Two major tasks were undertaken on the third school visit:

[10] After initial analysis of the analytical scoring of particular facets, it was not pursued as it did not seem to offer much information, and the facet descriptors implemented as a set produced satisfactory agreement.

- the overall assessor worked through the process of synthesising the teachers' assessments into one overall assessment; and
- the researchers distributed evaluation questionnaires and interviewed teachers about the process.

Overall Assessment

Just prior to this third visit, each Overall Assessor was sent a document which outlined how they might best deal with the material that would be presented during the visit. The Overall Assessor was expected to work through all of the assessments for that school noting, in particular, cases where the assessments given by teachers ranged over four or more points of the eight-point scale (eg. from 1 to 4). It was expected to be easy in some cases of discrepant marks to contract to an acceptable range. For example in cases where outlying teachers had indicated (by means of such symbols as ?, ↑ or ↓) that they would be prepared to have their marks shifted in a certain direction. Other cases would be more difficult and it was expected that there might be some cases that would not be resolved.

The object of this activity was to see whether, and to what extent, the Overall Assessors could and would be prepared to use particular pieces of information in making an overall decision about the most appropriate result for an individual student.

Overall Assessors seemed to have less difficulty with this exercise than had been anticipated. There were some unresolvable cases, although sometimes this was because the exercise had to be completed on the day of the visit and there was not always time to hold discussions with teachers whose marks for particular students had been discrepant.

Questionnaire and Interviews

A list was sent to schools which outlined the kinds of issues to be discussed with participants on the third visit. Most schools went to some trouble to ensure that participants could be available to take part in interviews at the busiest time of year.

Synthesis by the Overall Assessor

The project aimed to develop a prototype computer program that could be used for four interrelated purposes:

- to enable teachers to enter their KC assessments for each student;
- to enable the overall assessor to synthesise teachers' assessments into an overall level for each student;
- to produce the results on a report form; and

- to produce data about the assessments of use to the school.

The levels given and descriptors chosen by each teacher were merged into one file and the program was used to:

- calculate an outcome using a set of algorithms on each Key Competency for each student, taking into account the levels recorded by different teachers (see Appendix 6);
- indicate whether the range of levels is irreconcilable (and should be reported only after consultation with particular teachers);
- review the descriptors identified by individual teachers and determine the most appropriate descriptor for each student. The overall assessor reviewed the range of levels, the codes and comments given by teachers and the algorithmic outcome calculated by the program, to give each student a level for each Key Competency.

The Reporting of Levels

The program was used to produce reports indicating levels for each Key Competency and typical descriptors on each Key Competency for each student. The program was able to indicate if certain teachers were 'lenient' in their assessments and if others were 'harsh'. These data might be useful when the overall assessor is producing a level for a student.

Descriptive statistics were produced so that schools were able to compare assessments across the Key Competencies, across teachers and across subject areas.

Participants' Views of the Assessment Framework and Procedures

Participants were asked to fill in a questionnaire about the assessment framework and the assessment procedures and they were interviewed by a member of the project team. A detailed outline of the feedback is presented in Appendix 7.

Summary of Participants' Views of the Facet Descriptors

Overall, it seemed that the notion of Facets had worked for the participants and that the Facets developed were generally sufficiently, but not too, different from each other, so that they were perceived as defining a coherent concept. Teachers' responses indicated a remarkable degree of agreement and acceptance of the mapping of Key Competencies presented. For example, 90 per cent of the teachers who responded to the questionnaire indicated that the facets for Working in teams were 'different and coherent' - in other words the facets gave a, more or less, complete definition of the Key Competency but each expressed a different aspect of it. There was also a very high level of acceptance

about Planning and organising (84 per cent) and Solving problems (88 per cent). Some Key Competencies appeared to be a little more problematic, but closer inspection suggests that this was often due to the need for some refinement to the wording of Facets. This is the case with Communicating ideas and information - oral (66 per cent) and written (62 per cent) and Collecting, Analysing and organising (72 per cent). The conceptions of Using mathematics, Cultural understanding and Using technology require further thinking, although in all cases more than 50 per cent of the respondents indicated that the Facets were 'different and coherent'. Overall there was a significant degree of agreement that the definitions give to the Key Competencies were workable.

It is particularly interesting to note that the less cognitive or more attitudinal Key Competencies such as Planning and organising and Working in teams were so well received. In the past, educators have tended to shy away from formally assessing attitudes, yet it seemed that these Key Competencies operated even better than the Key Competencies that are more aligned with traditional curriculum areas - such as Using mathematics and written and oral communication.

Teachers found that the overall assessment framework was clear and that they were able to use it with remarkably little instruction. One teacher reported that it took only ten minutes to fully explain the procedure to a colleague who had been unable to attend the briefing session.

As might be expected, some teachers believe that the Key Competencies may provide an opportunity for less academically able students to achieve. A clear majority of teachers who completed the questionnaire believe that the Key Competencies would be useful for employers.

The Trial Results

Trial participants were interviewed after the assessment and they generally reported that they found the assessment framework was clear, that they were able to use it with little difficulty, and that they were able to make the global, impression judgements envisaged by the project.

The following discussion outlines some of the quantitative data obtained from the trial. The data collection and the analysis were designed to address two basic issues. The first issue is the characteristics of the outcomes produced by the assessment procedures, and the second is the reliability, in terms of the inter-rater agreement or consistency, of the assessments made by the participating teachers.

The analysis reported here deals with the following questions:

- How did the participating teachers use the Key Competencies score scale?
- What sorts of results did they produce?
- Did the participants discriminate among the students?
- How much did the participants agree with each other in the way they used the score scales and the way they assessed individual students?

Tables 31 to 33 below give an indication of the properties of the scores produced by the assessment trial. Tables 34 to 36 give an indication of the agreement between the assessing teachers.

Use of the Score Range

Table 31 shows that the scores given by all teachers for each competency are more or less normally distributed with 5 being the clear modal mark in all cases.

The cumulative percentages at and below the modal mark in Table 31 give an indication of the view of the assessors on each competency. The KC3 Mathematics assessments were the lowest range of scores with 75% at or below 5. KC6 Technology assessments are also lower with 74% at or below 5. The KC8 Teamwork assessments are the highest range of scores with only 63% at or below 5. KC2a Oral (at 65%) KC5 Problems solving (at 66%) are also higher.

Figures 17 to 25 give histograms of the percentage for each score point given by all participants for all Key Competencies. In general these histograms show that teachers discriminated quite significantly among the students. Although comparatively small numbers of students were given scores of 1, 2 and 8, most students were spread between 3 and 7, effectively range of five points.

As was expected, the division of three levels into two sub-levels produced the 'saw-toothed' distribution to be seen in the histograms. In general scores of 3, 5 and 7 were more commonly given than 4 and 6, and in all cases 5 was the modal mark.

Apart from the dips at 4 and 6, the distribution shown in the histograms is approximately normal, as would be expected. The dips result from the presentation of the scale as having 3 levels and each divided into 2 sub-levels. This has the impact of encouraging assessors to place students just into a level, that is place them at the bottom sub-level. If the assessors were tempted to give the student the top of the level, they were also tempted to move the student into the next level, and hence the dips at scores 4 and 6.

The histograms show that teachers did not inappropriately bunch the students and that they were making meaningful discriminations among them.

With the exceptions of KC3 Maths, KC6 Technology and K4 Cultural understandings, teachers were able to produce assessments of most of the Key Competencies. Assessments of KC3 Maths were not produced by teachers in 47% of cases. In 30% of cases assessments of KC 6 Using technology were not produced. KC4 Cultural understandings was not assessed for 15% of cases. The number of assessments and the distributions of the scores were quite consistent for the other Key Competencies.

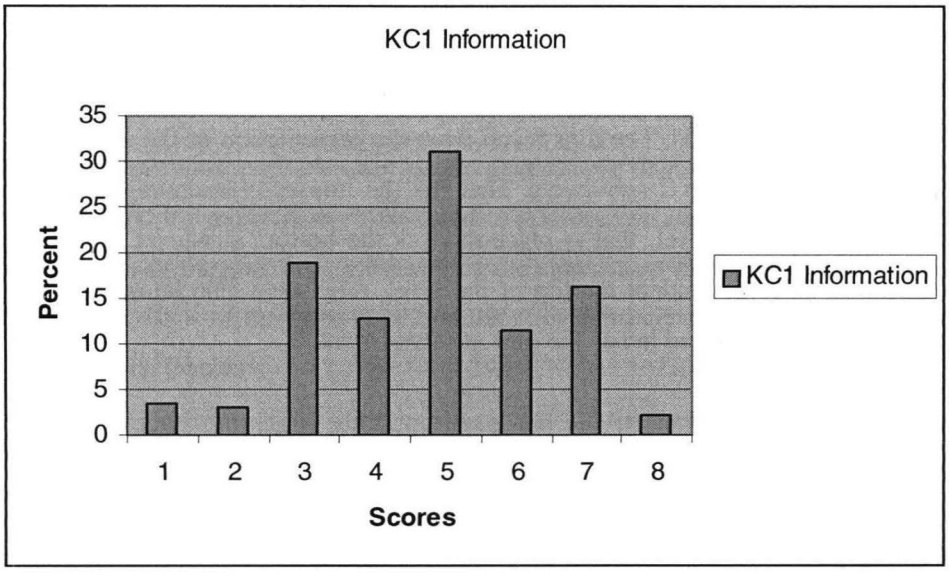


Figure 17 The Distribution of Scores for Collecting and analysing information

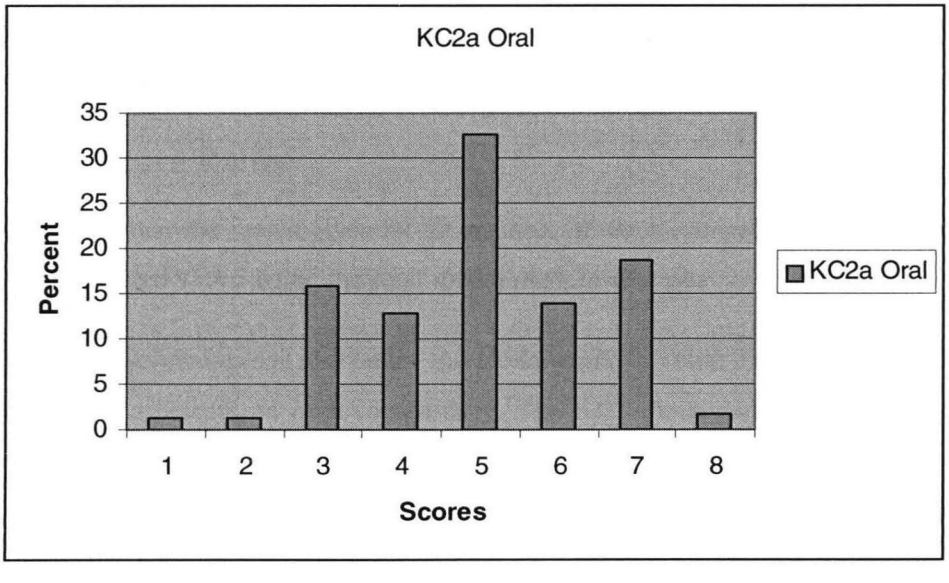


Figure 18 The Distribution of Scores for Oral communication

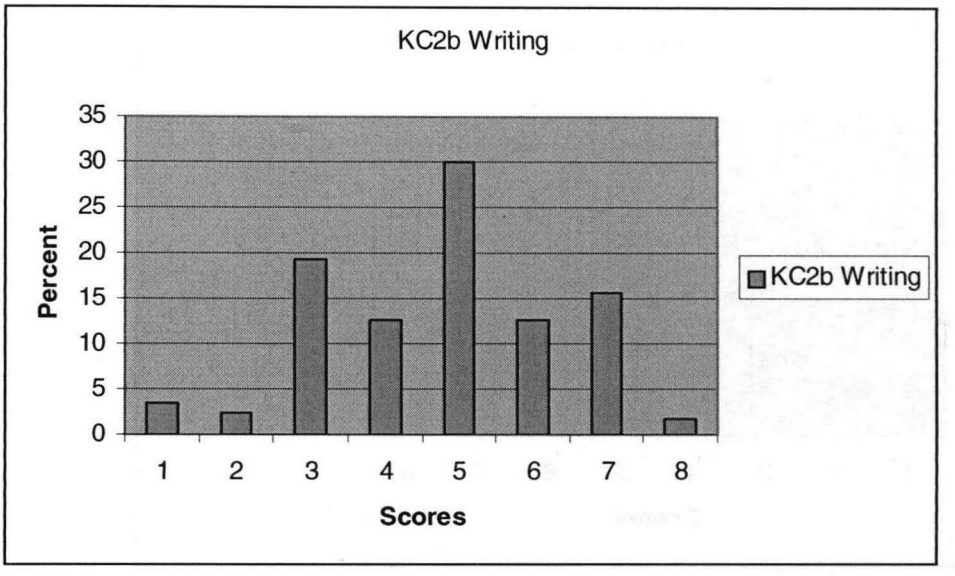


Figure 19 The Distribution of Scores for Written communication

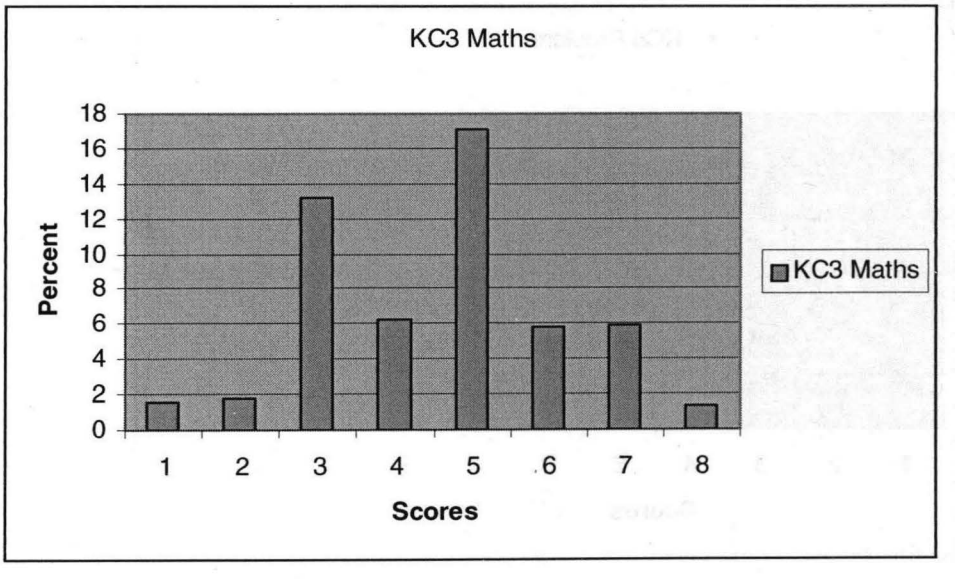


Figure 20 The Distribution of Scores for Using mathematics

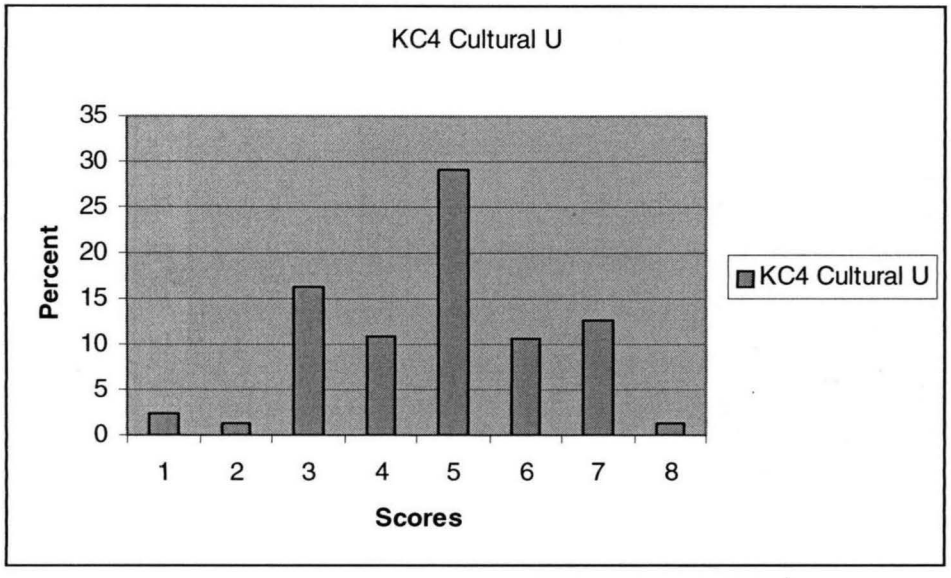


Figure 21 The Distribution of Scores for Cultural understanding

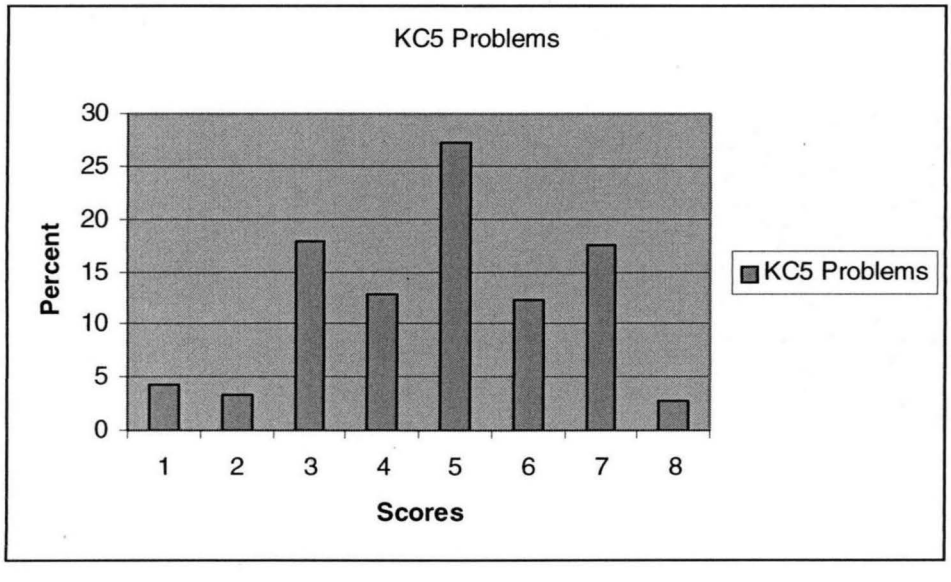


Figure 22 The Distribution of Scores for Problem solving

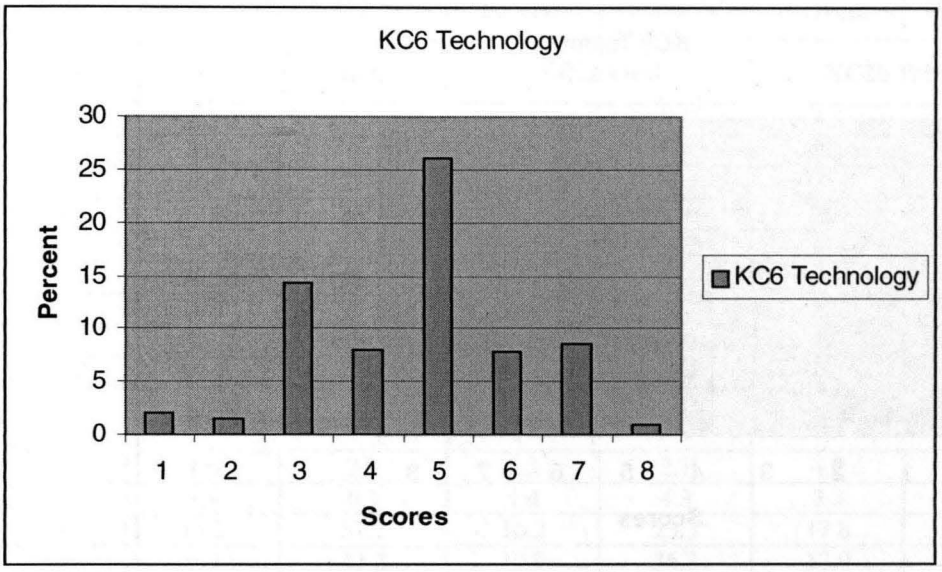


Figure 23 The Distribution of Scores for Using technology

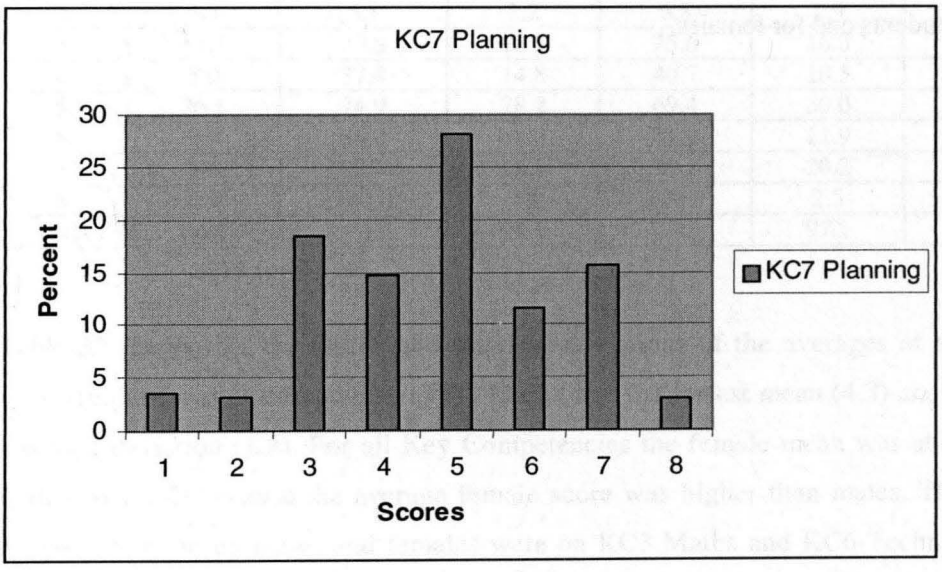


Figure 24 The Distribution of Scores for Planning and organising activities

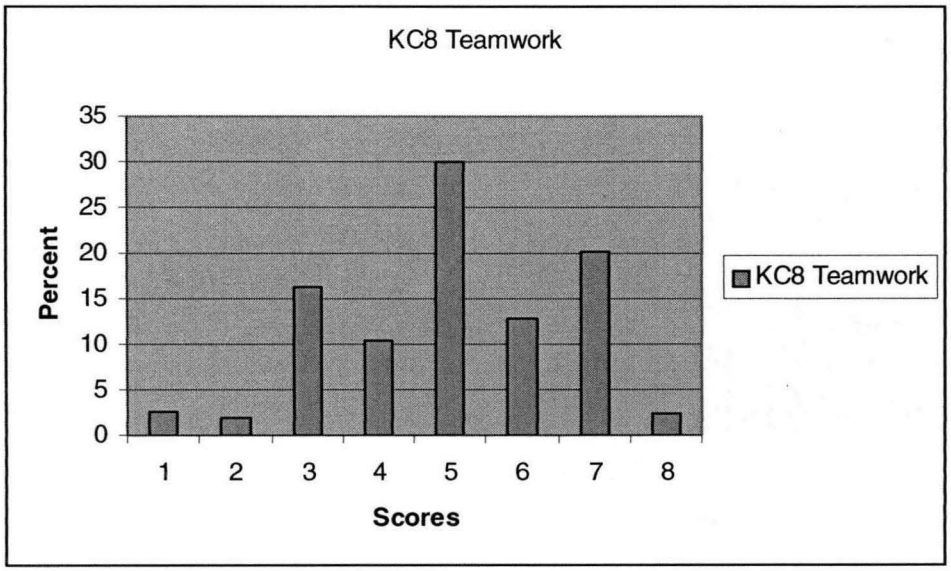


Figure 25 The Distribution of Scores for Working with others and in teams

Other than the individual marks of teachers shown in Table 31, it is useful to average the scores given by the different teachers for the same student. Tables 32 and 33 and Figure 26 are based on the average of all marks for each student, and Tables 32 and 33 also include the results all students and for females.

Table 31 Frequencies of All Assessment for All Competencies

Score	Percent	Cum. %	Percent	Cum. %	Percent	Cum. %
	KC1 Information		**KC2a Oral**		**KC2b Writing**	
1	3.5	3.5	1.4	1.5	3.5	3.6
2	3.0	6.6	1.4	2.9	2.3	5.9
3	18.9	25.7	15.8	19.0	19.3	25.6
4	12.8	38.7	12.8	32.0	12.7	38.6
5	31.1	70.0	32.7	65.2	30.0	69.3
6	11.5	81.6	14.0	79.4	12.7	82.2
7	16.2	97.9	18.6	98.3	15.6	98.1
8	2.1	100.0	1.7	100.0	1.8	100.0
Total	99.1		98.4		98.0	
	KC3 Mathematics		**KC4 Cultural U**		**KC5 Solving Probs**	
1	1.6	2.9	2.3	2.7	4.2	4.3
2	1.8	6.3	1.4	4.3	3.3	7.7
3	13.2	31.2	16.2	23.5	17.8	25.8
4	6.3	43.2	10.9	36.3	12.9	39.0
5	17.1	75.5	29.2	70.8	27.2	66.7
6	5.8	86.4	10.7	83.4	12.3	79.2
7	5.9	97.6	12.6	98.3	17.6	97.1
8	1.3	100.0	1.4	100.0	2.8	100.0
Total	52.8		84.7		98.2	
	KC6 Technology		**KC7 Planning**		**KC8 Teamwork**	
1	2.1	3.0	3.5	3.6	2.7	2.8
2	1.4	5.1	3.2	6.8	1.9	4.7
3	14.3	25.8	18.5	25.6	16.3	21.6
4	8.0	37.4	14.8	40.7	10.5	32.3
5	26.1	74.9	28.2	69.4	30.0	63.3
6	7.9	86.3	11.6	81.2	12.9	76.6
7	8.6	98.7	15.6	97.1	20.2	97.4
8	0.9	100.0	2.9	100.0	2.5	100.0
Total	69.4		98.1		97.1	

Table 32 shows that the means and standard deviations of the averages of all students were similar for each competency. KC3 Maths was the lowest mean (4.3) and the largest standard deviation (1.8). For all Key Competencies the female mean was above that of both sexes (all) because the average female score was higher than males. The smallest differences between males and females were on KC3 Maths and KC6 Technology. The differences between female and males averages that can be inferred from Table 32 are not as substantial as the difference for both sexes between competencies (4.3 for KC3 and 5.1 for KC8).

Table 32 Mean and Standard Deviation of Average Scores of All and Females

	MEAN KC1	MEAN KC2A	MEAN KC2B	MEAN KC3	MEAN KC4
Mean all	4.8	5	4.8	4.3	4.8
Std Dev all	1.4	1.2	1.4	1.8	1.3
Female M	5	5.3	5.1	4.4	5
Female SD	1.3	1.2	1.3	1.7	1.2
Diff of all & female	-0.2	-0.3	-0.3	-0.1	-0.2
	MEAN KC5	**MEAN KC6**	**MEAN KC7**	**MEAN KC8**	
Mean all	4.8	4.6	4.8	5.1	
Std Dev all	1.4	1.5	1.4	1.3	
Female M	5.1	4.7	5.1	5.4	
Female SD	1.3	1.4	1.3	1.2	
Diff of all & female	-0.3	-0.1	-0.3	-0.3	

Table 33 shows that 5 was the modal score for the average for each competency, although for KC5 Problem solving and KC8 Teamwork an average score of 6 was as common or almost as common as a score of 5. The KC3 Maths assessments were the lowest range of scores with 75% at or below 5. KC6 Technology was also lower at 74% at or below 5. The KC8 Teamwork was the highest range of scores with only 59% at or below 5. KC2a Oral (at 63%) KC5 Problems solving (at 65%) were also higher.

Table 33 Distribution of Average Scores for Each Candidate

	KC1 Information			KC2a Oral			KC2b Writing		
	Freq	%	Cumul	Freq	%	Cumul	Freq	%	Cumul
1	6	1.0	1.0	1.0	0.0	0.0	4.0	1.0	1.0
2	20	4.0	5.0	8.0	2.0	2.0	20.0	4.0	5.0
3	56	11.0	16.0	43.0	8.0	10.0	64.0	13.0	17.0
4	117	23.0	39.0	112.0	22.0	32.0	100.0	20.0	37.0
5	157	31.0	70.0	156.0	31.0	63.0	158.0	31.0	68.0
6	97	19.0	89.0	133.0	26.0	89.0	103.0	20.0	88.0
7	47	9.0	98.0	51.0	10.0	99.0	54.0	11.0	99.0
8	8	2.0	100.0	4.0	1.0	100.0	5.0	1.0	100.0
	KC3 Mathematics			**KC4 Cultural U**			**KC5 Solving Problems**		
0	30	6.0	6.0	4.0	1.0	1.0			
1	12	2.0	8.0	5.0	1.0	2.0	5.0	1.0	1.0
2	12	2.0	11.0	8.0	2.0	3.0	23.0	5.0	6.0
3	100	20.0	30.0	56.0	11.0	14.0	61.0	12.0	18.0
4	99	19.0	50.0	111.0	22.0	36.0	109.0	21.0	39.0
5	129	25.0	75.0	172.0	34.0	70.0	132.0	26.0	65.0
6	84	17.0	92.0	106.0	21.0	91.0	127.0	25.0	90.0
7	37	7.0	99.0	44.0	9.0	100.0	47.0	9.0	99.0
8	5	1.0	100.0	2.0	0.0	100.0	4.0	1.0	100.0
	KC6 Technology U			**KC7 Planning**			**KC8 Teamwork**		
0	12.0	2.0	2.0	1.0	0.0	0.0			
1	7.0	1.0	4.0	5.0	1.0	1.0	4.0	1.0	1.0
2	13.0	3.0	6.0	21.0	4.0	5.0	13.0	3.0	3.0
3	67.0	13.0	19.0	61.0	12.0	17.0	45.0	9.0	12.0
4	114.0	22.0	42.0	110.0	22.0	39.0	103.0	20.0	32.0
5	162.0	32.0	74.0	149.0	29.0	68.0	135.0	27.0	59.0
6	96.0	19.0	93.0	111.0	22.0	90.0	139.0	27.0	86.0
7	35.0	7.0	100.0	42.0	8.0	98.0	66.0	13.0	99.0
8	2.0	0.0	100.0	8.0	2.0	100.0	3.0	1.0	100.0

Figure 26 shows the distribution of average scores. 5 was the clear modal score in all but KC8. The distributions of scores were fairly symmetrical in all cases, although KC3 has as many scores of 3 as 4, KC5 has almost as many scores of 6 as 5, and KC8 has more scores of 6 than 5.

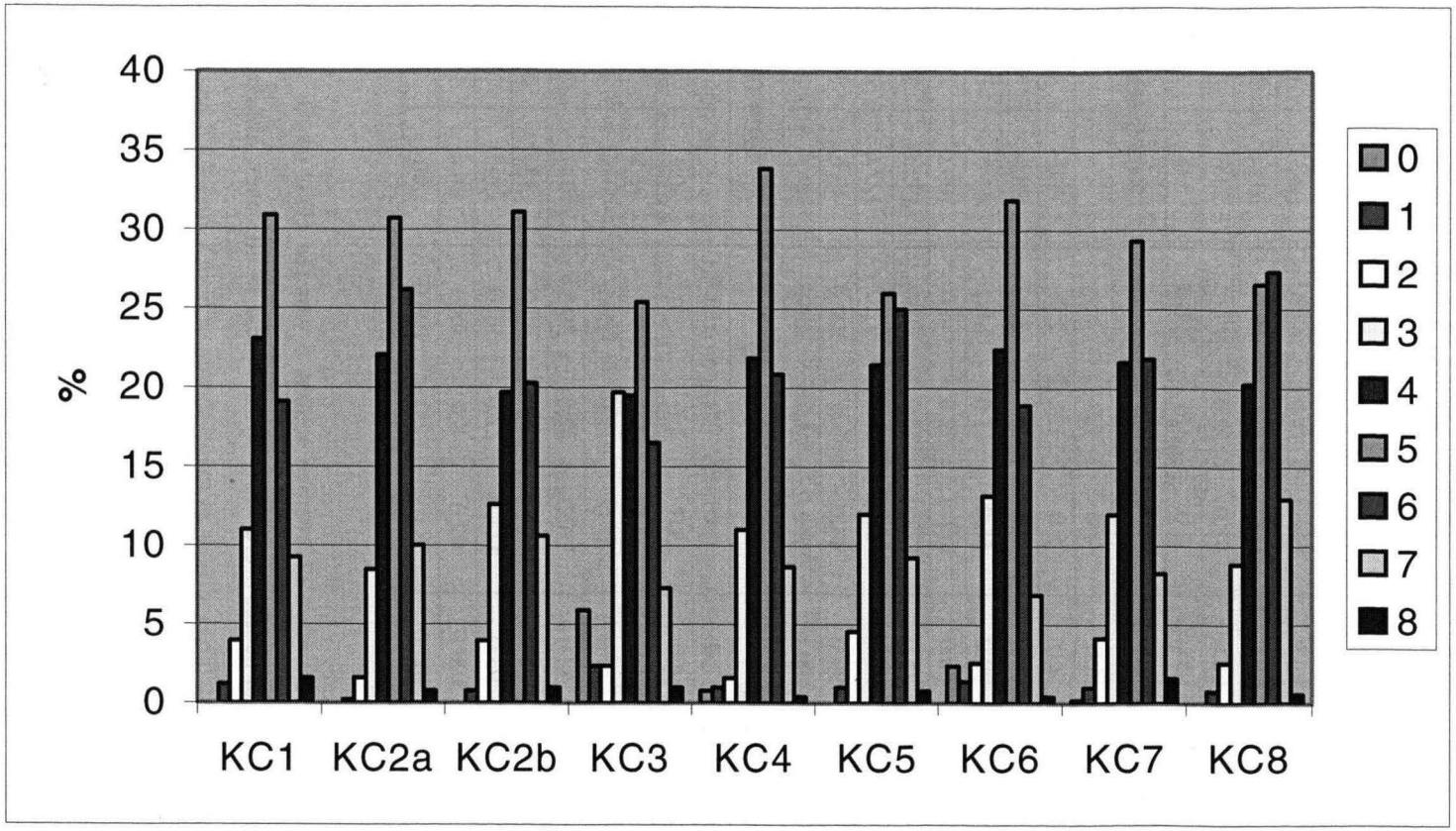


Figure 26 Distribution of Average Scores for Each Key Competency

In summary, Table 31 shows that individual teachers used the score range and discriminated quite significantly amongst the students. Tables 32 and 33 show that the averages of the scores also discriminated quite effectively among the students.

While it is important that the teachers, individually and as a group, discriminate amongst the students, it is also important that the teachers are more or less consistent in the way they see individual students.

Inter-rater Agreement

Table 34 shows the degree of agreement between different teachers assessing the same student. A difference of three score points (1 to 4, 5 to 8 etc.) on an 8 or 10 point scale is commonly judged to be unacceptably discrepant in a structured marking procedure dealing with a single piece of work. Marking procedures usually require some kind of resolution of this discrepancy. Such two marker discrepancies in controlled marking of a single piece of work should be under 10% and are good if they are under 5% on an 8 to 10 point scale.

Rows 0 to 6 of Table 34 show the differences between the scores of pairs of students assessing the same student. Row 0 is where there was no difference between two teachers, row 1 is where there is a difference of 1, and so on. A difference of 1 is the difference between scores of 3 and 4 or 7 and 8. A difference of 2 is the difference between scores of 3 and 5 or 6 and 8. The last row for each KC of the table shows the discrepancies of more than 2 points difference. The average for adjacent grades was 55.8%, and the average for grades separated by a difference of 1 or 2 is 88.8% in these pairwise comparisons. These figures give an average pairwise discrepancy rate of 11.2% for these global impression judgements.

Table 34 Percentage Agreement in Pairwise Comparison

	KC1 Information			KC2a Oral			KC2b Writing		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	668	31.4	31.4	598	28.6	28.6	658	31.8	31.8
1	701	33.0	64.4	711	34.0	62.6	688	33.2	65.0
2	544	25.6	90.0	573	27.4	90.0	510	24.6	89.7
3	146	6.9	96.8	136	6.5	96.5	149	7.2	96.9
4	55	2.6	99.4	70	3.3	99.8	57	2.8	99.6
5	9	0.4	99.9	4	0.2	100.0	5	0.2	99.9
6	3	0.1	100.0				3	0.1	100
Total	2126			2092		100	2070		
Disc		10.0			10.0			10.3	
	KC3 Mathematics			**KC4 Cultural U**			**KC5 Solving Problems**		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	168	29.6	29.6	454	28.7	28.7	571	27.4	27.4
1	188	33.2	62.8	540	34.1	62.8	676	32.5	59.9
2	156	27.5	90.3	418	26.4	89.3	538	25.8	85.7
3	45	7.9	98.2	103	6.5	95.8	188	9.0	94.7
4	9	1.6	99.8	54	3.4	99.2	85	4.1	98.8
5	1	0.2	100.0	8	0.5	99.7	20	1.0	99.8
6				5	0.3	100	5	0.2	100
Total	567			1582	100		2083		
Disc		9.7			10.7			14.3	
	KC6 Technology U			**KC7 Planning**			**KC8 Teamwork**		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	298	29.9	29.9	585	28.2	28.2	546	27.2	27.2
1	300	30.1	60.0	712	34.3	62.5	676	33.6	60.8
2	298	29.9	90.0	542	26.1	88.6	555	27.6	88.4
3	68	6.8	96.8	162	7.8	96.4	141	7.0	95.4
4	26	2.6	99.4	66	3.2	99.6	74	3.7	99.1
5	2	0.2	99.6	8	0.4	100.0	14	0.7	99.8
6	4	0.4	100.0	1	0.0	100.0	5	0.2	100.0
Total	996			2076			2011		
Disc		10.0			11.4			11.6	
Disc average	10.9%								

Table 35 is a summary of the 2 and 3 point differences from Table 34.

Table 35 Percentage of Students Given a Difference of Three or Less by Pairs of Teachers

	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
% of no difference	31.4	28.6	31.8	29.6	28.7	27.4	29.9	28.2	27.2
% at or below a difference of 1	64.4	62.6	65.0	62.8	62.8	59.9	60.0	62.5	60.8
% at or below a difference of 2	90.0	90.0	89.7	90.3	89.3	85.7	90.0	88.6	88.4
Average difference of 1 or less	62.3%								
Average difference of 2 or less	89.1%								

Table 36 gives a comparison to assist the interpretation of the results in Tables 34 and 35. It shows the results of pairwise comparisons between the trained markers of a single piece of work in a supervised marking regime assessing (on a 10 point scale to score writing ability) of Year 12 students. Table 36 shows that the global, impression judgements of the Key Competencies in this trial are not unlike the results from this supervised test-marking arrangement. The pairwise comparison for the Key Competencies grading is 89.1% with a difference of 2 or less grades on average and the pairwise comparison for the writing test is 94.3% with a difference of 2 or less.

Table 36 Results of Writing Test Marking on a Ten Point Scale

Pairwise comparison of writing test markers on 10 point scale		
Mark range	**Percent**	**Cumulative %**
0	29.0	29.0
1	44.5	73.5
2	20.5	94.3
3	5.5	99.8
4	0.5	100.0

Validity and Soft-edged Reliability

Validity and reliability are key terms in the discourse of assessment.

Reliability is the consistency and accuracy of an assessment. A reliable assessment will yield similar results over time with similar populations in similar circumstances. A reliable assessment can be generalised. The kind of reliability at issue in whole-school assessment is inter-rater reliability. The reliability of whole-school assessment is a matter of the extent to which different assessors agree with each other in assessing individuals and groups of students.

The validity of an assessment is the extent to which it assesses what it intends or claims to assess. Validity is a very complex concept that has many permutations. Construct validity is concerned with the definition of dimensions or criteria. Predictive validity is concerned with the ability of an assessment to predict future performance. Concurrent validity is concerned with the extent to which an assessment agrees with other assessments of the same construct. Convergent and discriminant validity is concerned with the extent to which a set of assessments (as in Table 37) form an intelligible and interpretable pattern of high and low correlations.

It is sometimes argued that there is a potential conflict between validity and reliability. A test of grammatical knowledge, for instance, may be reliable in that it consistently discriminates among candidates, but if the test is claimed to be an indirect measure of writing ability it may be of limited validity.

In writing about the assessment of the Key Competencies in 1996 and discussing the data that might result from whole-school assessment I have used the term 'soft-edged reliability'. This is a rhetorical device which I used (and still find useful) to avoid simple judgements of 'more and less reliable'.

In analysing assessment options for the Key Competencies I argued that an assessment based on multiple teacher judgements would not readily attain the kind of precise or 'hard-edged reliability' needed for comparing different schools or systems. In such assessments small differences need to be and are significant. On the other hand whole-school assessment does not require a hard edge. Whole-school assessment can be broad, general assessments but still carry meaningful information.

In my view it is not useful to say an assessment like the Key Skills test developed by the Qualifications and Curriculum Authority in Britain is very reliable. Whereas a group of teacher judgements is less reliable. By using the term soft-edged I am distinguishing the kind of reliability attained by synthesizing teacher judgements from that attained by a multiple choice test.

The whole-school assessment that I have developed is soft edged in more than name. As will be seen in the next chapter, soft-edged assessment does not produce a result which is a single point on a scale. The results of whole-school assessment are reported as a region of one, two or three points on the eight point scale. The outcomes of whole-school assessment are soft-edged in that they can be broader or narrower depending on the degree of agreement between the different teachers of a student.

Conclusions from the SBKCLAP

In a number of respects the results of this trial are surprising and remarkable, and they are at odds with the claims made by critics of the Key Competencies that:

- student performance is significantly different in different subjects;
- different teachers will have quite different perspectives on students; and
- it is not possible or meaningful to generalise about students across the curriculum.

The results of the trial show a significant degree of agreement in their assessments of the same student between teachers of different parts of the curriculum. The agreement is comparable with the agreement between pairs of markers in a supervised assessment of a single piece of writing. The trial results suggest that such whole-school and cross-curricular assessments can be psychometrically sound, and that they can carry useful information. The trial shows that teachers can distinguish what they take to be general about a student from what they take to be subject-specific or particular to their relationship with the student. The trial also shows that teachers see and know a good deal more about students than is included in subject assessments, and that this understanding can be drawn on in a cross-curricular assessment. The trial results suggests that the performance of students across different subject areas shows less variation than we might expect or assume.

The teachers within a school as a group seem to discriminate consistently among their students, and this trial also suggests that the Key Competencies can be usefully assessed at provider level without a large investment of time or resources. The procedures developed in this trial were judged by participants to be efficient and cost-effective. Teachers were able to make assessments of a student in two or three minutes using the method developed in the trial, and the procedures give the basis for producing integrated, whole-schools assessments efficiently.

The outcomes of this trial were subsequently confirmed in much less elaborate trials with 10 Victorian schools in 2001, 15 schools in 2002, 20 schools in 2003 and 40 schools in 2004.

Limitations of Whole-school Assessment

The major psychometric limitation of the data generated by this trial is found in the correlations between the competencies presented in Table 37. Some of the correlations between different Key Competencies are in the low 0.5s, others are in the high 0.7s to 0.8s.

These correlations do not offer the kind of discriminant and convergent pattern of a really satisfactory differential ability construct. One would wish to see a clearer pattern of separation between individual Key Competencies and different clusters of Key Competencies.

It might be argued that the data show a general ability underpinning performance as a whole. Or it might be argued that the Key Competencies are not well framed as a differential ability theory, and that they do not maximally focus on the most important patterns of different strengths and weaknesses in students. Or it may be that the global impression judgements made in this assessment operate so that a 'halo effect' makes the different judgements similar rather than differentiated. Whether the latter is the case could be tested by designing a set of generic skills that aimed to better meet psychometric standards of convergent and discriminant validity.

The assessments undertaken in this trial were quite modestly supported. Participants were asked to use an unfamiliar framework with little guidance. Participants were told that they should be able to make an assessment of students in two or three minutes reflection. The intention was not to finely nuance the particular set of constructs, but rather to explore the degree of agreement between different teachers in rather general terms. Whatever the explanation, the degree of positive manifold to be seen in these Key Competencies assessments reported here suggests that further work on the definition of the Key Competencies and the assessment procedures will need to aim for more significant differentiation between competencies. The agreement between

teachers in the trial was significant. Perhaps a whole-school assessment procedure that was built into a developed program would produce a subtler set of assessments of individuals than was achieved in this study.

Potential Uses of Whole-school Assessment

The crucial aspect of the whole-school assessment procedures trialled here is that they aim and manage to produce an overall or consensus judgement of generic, cross-curricular skills in an efficient and cost-effective fashion.

The second crucial aspect of this method of whole-school assessment is that it is able to deal with the difficult issues of assessing attitudes and personal growth because it is a whole-school position rather than the judgement of any individual. These procedures offer schools the opportunity to monitor attitudes and personal growth (as well as more academic matters) in a collaborative but cost-effective fashion.

Table 37 Inter-correlations of Key Competency Judgements

	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8	Range
KC1		.67	.83	.78	.71	.77	.64	.75	.65	.19
KC2a			.70	.59	.67	.65	.55	.65	.69	.28
KC2b				.74	.70	.75	.62	.74	.65	.13
KC3					.52	.72	.67	.67	.57	.20
KC4						.69	.53	.70	.65	.24
KC5							.69	.83	.72	.18
KC6								.68	.58	.15
KC7									.73	.16
KC8										

10

The Whole-school Assessment Process and Software

The previous chapter reported on the initial trial of whole-school assessment made in 1996. Reference has been made to the further work on whole-school assessment in the National Industry Education Forum's Key Competencies Portfolio Project and the ongoing work for the Victorian Board of Studies and the Victorian Curriculum and Assessment Authority. In this more recent work the basic outline of whole-school assessment has remained much as it was when first developed in 1996. Even so the whole-school assessment process has been refined over the last six years. These developments and the process itself can be best understood by examining the Whole-school Assessment Software I have designed and used in the Key Competencies Portfolio Project and the Key Competencies assessment projects for the VCAA.

An Overview of the WSA Process and Software

The Whole School Assessment (WSA) process is based on assessments made by a range of teachers of students in a school so as to produce an overall school assessment and report. The process is supported by a database for the recording, integration and reporting of teacher assessments of student performance. The form of the student report produced by the software is shown in Figure 27.

In this assessment process, different teachers assess the same things while teaching different subjects. They make these assessments by drawing on the knowledge they gain of students in a range of different school activities. The aim of this assessment process is to produce a general or whole-school report rather than a number of separate reports by individual teachers within different subject areas. This WSA process contrasts with most other school assessments because it is not subject-based. Unlike subject-based assessments, the WSA process is general or cross curricular, and it can also take into account extra curricular and other activities. The object of the process is to produce a whole-school or common view of the student in a single, integrated report.

Report for Doug McCurry

July 2004 **Grade 12**

Teacher Judgements of Levels of Performance

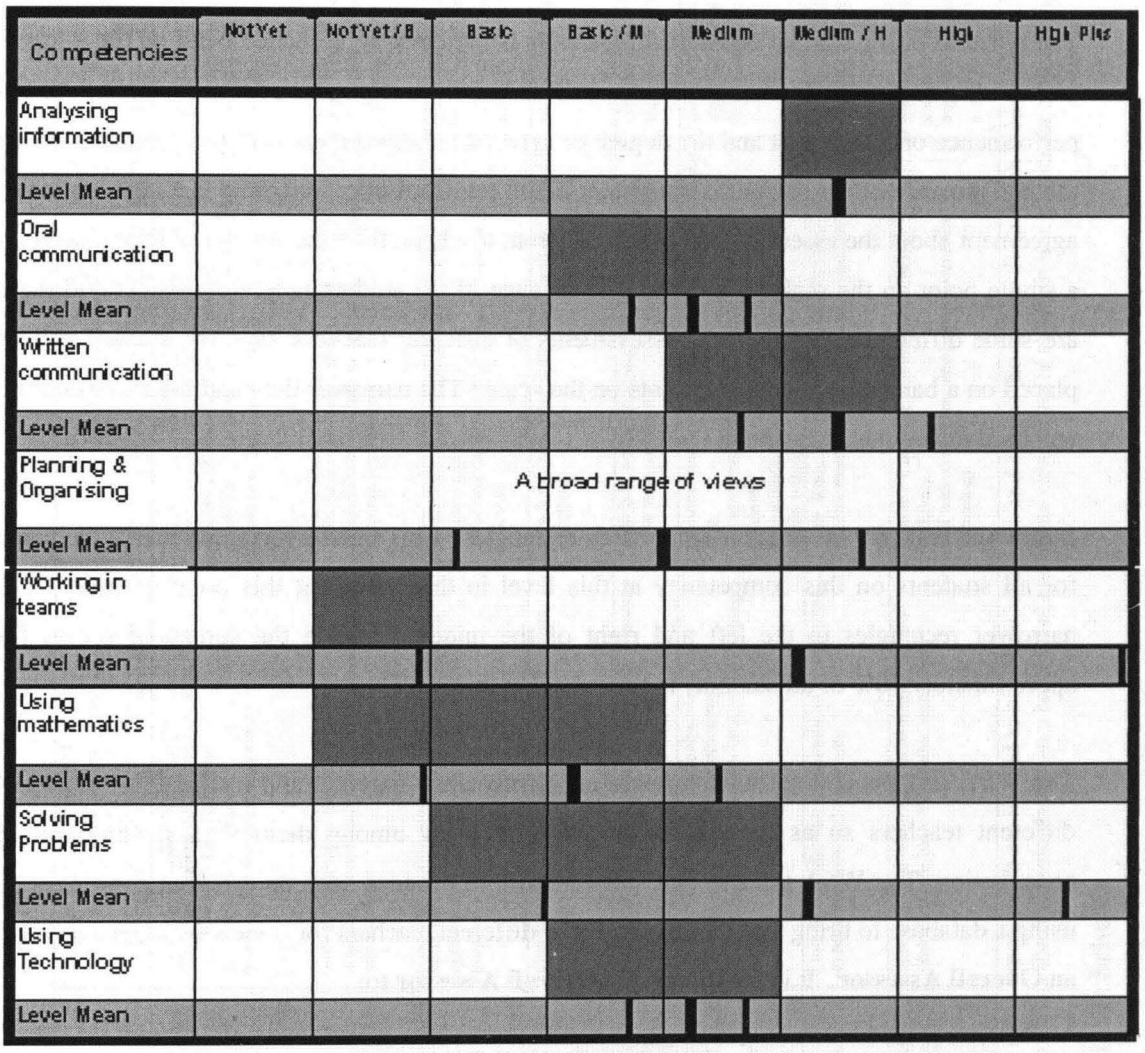


Figure 27 The WSA Software Report Format

The WSA software reporting format

Figure 27 shows the overall judgement determined by a school for a student on each assessed competency. These judgements are based on a range of levels from Not Yet level 1, through Levels 1 and 2 to Level 3, the top level. These levels are further divided into an 8 point scale. These levels and the associated verbal descriptors are shown in the table 38.

The position of the student on each competency is indicated by a shaded band on the 4 levels and the 8 point scale. The range for the student can be wider or narrower depending on the performance of the student and the degree of agreement between the different teachers. When the performance of the student is consistent in different activities and there is a high degree of agreement about the assessment between different teachers, then the student will be placed on a single point on the scale. When the performance of the student is less consistent and there are some differences between the assessments of different teachers, then the student will be placed on a band of two or three points on the scale. The narrower the band the more definite and hard-edged the assessment.

Below the band for the student are 3 black rectangles. This middle rectangle marks the mean for all students on this competency at this level in this school at this point of time. The narrower rectangles to the left and right of the middle indicate the spread of scores for approximately 70% of all students in the school at this level.

The WSA process and software aims to efficiently draw together and integrate the views of different teachers so as to produce a consensus view among them with a minimum of consultation. The WSA process and software aim to produce such an efficient consensus by using a database to bring together the views of different teachers for review by a person called an **Overall Assessor.** It is the role of the **Overall Assessor** to:

- review the outcomes suggested by the software and the assessments of different teachers;
- make decisions in some cases about the consensus view to be reported; and
- consult with some or all teachers to find a consensus view in discrepant cases.

This role and the technology supporting it are at the centre of the WSA process.

Levels for Grades 9 & 10	Facet Descriptors	Level 1 Basic achievement		Level 2 Medium achievement		Level 3 High achievement	
Codes	Aspects of the Competency	B 3	B/M 4	M 5	M/H 6	H 7	H+ 8
		The student is able to		The student is able to		The student is able to	
1 Collecting, analysing & organising ideas & information	locating and collecting ideas & information comprehending and interpreting ideas & information analysing and evaluating ideas & information	deal with basic and routine research tasks comprehend basic information organise basic information assess basic information as directed		deal with some more complicated research tasks comprehend & analyse basic information organise some more complicated information. assess basic info. independently		deal with complicated research tasks comprehend and analyse more complicated information organise more complicated info. assess	
2a Written Communication of ideas & information	writing accurately and conventionally writing clearly and coherently using formal and informal styles appropriately	write on some topics clearly and coherently appreciate the demands of writing for some purposes and audiences		write more complicated ideas clearly and coherently appreciate the demands of writing for a range of purposes & audiences		write complicated ideas clearly & coherently appreciate the demands of writing for a wide range of purposes and audiences	
2b Oral Communication of Ideas & Information	speaking clearly and precisely using formal and informal speaking styles appropriately responding in oral interchanges	speak precisely and appropriately on some topics appreciate the demands of speaking for some purposes and audiences		speak precisely and appropriately in a range of ways on some topics appreciate the demands of speaking for a range of audiences		speak precisely and appropriately in a wide range of ways on a range of topics appreciate the demands of speaking for a wide range of purposes and audiences	
3.Planning & Organising Activities	thinks ahead and anticipates possible problems is systematic and practical seeks organisational challenges and experiences	plan & organise some basic and defined activities		plan & organise some more complicated and defined activities		initiate, plan & organise complicated activities	
4 Working with Others & in Teams	adapts to and contributes to group processes is sensitive to and supportive of others in a group seeks opportunities to work in teams and with other	take a role in a team and support other team members		contribute well to team work and usefully support other team members		contribute significantly to team work and effectively support other team members	
5. Using Mathematical Ideas & Techniques	understands mathematical principles and procedures selectively and appropriately uses mathematical techniques applies mathematical knowledge to solve problems	use basic maths methods to analyse data use basic maths reliably and efficiently appropriately select and apply basic maths techniques		use some more complicated maths methods to analyse data use some more complicated maths reliably and efficiently appropriately select and apply more complicated maths techniques		use complicated maths methods to analyse data use complicated maths reliably and efficiently design, interpret and evaluate maths activities	
6 Solving Problems	can see problems as opportunities responds positively and persistently to the challenge of problems seeks opportunities to solve problems	show focus, independence, &initiative in responding to basic problems and challenges		show focus, independence, & initiative in responding to some more complicated problems and challenges		anticipate problems & show independence & initiative in dealing with complicated problems & challenges	
7 Using Information Technology	can understand information technology systems is interested in underpinning principles as well as uses of information technology seeks information technology experiences and challenges	approach some information technology positively deal with some basic information technology		approach a range of informaiton technology positively deal with some more complicated information technology		approach a range of information technology with energy and application deal with more complicated information technology	

In this assessment, students who do not meet basic expectations are judged to be below level one, and those who clearly distinguish themselves as high level performers are at level three. The distinction between Levels 1 and 2 is the difference between a basic level and some clear signs of strength in a particular competency.

Table 38 The WSA Facets and Standards Schema

Global Impression Judgements of Generic Skills

The judgements made in the WSA process are concerned with that which is general to education (and might be properly called cross-curricular competencies) rather than the knowledge and skills that are more or less specific to individual subjects and subject-based assessments. The WSA assessment process might be described as being based on contextualised and global impression judgements made by individual teachers that are synthesised into overall school assessments.

If the competencies assessed in the WSA process are to be general, they must take into account as much of the school activity of a student as possible. In practical terms, an overall judgement that arises from a range of teacher perspectives can be seen as escaping the limitations of any individual teacher's perspective and, as a result, increasing the reliability of the assessment.

The judgements made by individual teachers in the WSA process are 'contextualised' in that they are made, primarily, on the basis of what teachers see in the course of teaching and assessing their normal courses.

The judgements teachers are asked to make might be described as 'global impressions' because they are not based on specific tasks or instances of behaviour. They are to be overall and synthesised impressions that can take into account as much evidence as the teacher can register. The assumption behind this assessment is that teachers do gain, to varying degrees, insight into the general competency achievements of their students, and this assessment aims to draw on and formalise these impressions.

It is crucial that the competency judgements made by individual subject teachers in the WSA process are general rather than subject specific. Teachers will, of course, gain most of the knowledge of students from subject classes, but teachers in this kind of assessment are not being asked to make judgements about that which is specific to their subject area. The WSA competency judgements are not about students' performances in particular subject areas. They are judgements about what the particular teacher takes to be general about the abilities or the performances of a student. These judgements are broad or global inferences about what is taken to be typical of the students in most subject areas, and other activities.

Teachers of different subject areas have different perspectives on students, and differences in their competencies judgements are to be expected. But such differences should not result from teachers thinking they are assessing different things or because they assume that what is true of their subject or their personal interaction with the student is typical or general.

For instance, a student who has a gift for music may be the leader of a school orchestra, but this does not mean that the music teacher who organises the orchestra should automatically assume or claim that such activity amounts to a high level of Working with others and in teams in general. The performance of such a student is the basis for claiming a level 3 competency if the music teacher considers it a general characteristic of the student. It is possible that although the music teacher observes that student working well with others in the orchestra, s/he is doubtful about the student's ability to work very well in teams in general, and so does not believe the student is typically at Level 3 in terms of team work. Other examples of the difference between subject-specific performances that might be recognised in subject grades and the typical or general performances that are assessed in WSA competency levels might easily be elaborated.

The crucial point for the WSA assessment is that these competencies judgements are about what the teacher infers from what s/he observes about the general rather than the subject specific. This point is crucial not only for theoretical reasons. If an overall competency judgement is to be practical, valid and reliable, different teachers should think they are aiming to do the same thing, that is, aiming to assess what is typical and general about the performance of students. If different teachers produce many cases of inconsistent and irreconcilable judgements about individual students, then the notion of an overall competencies judgement is invalid, as well as being unreliable and impractical.

It is presumed that contact in subject classes is the basis of this competencies assessment, but other information gathered from co- and extra-curricular activities and work placements can be taken into account.

Making teacher judgements of competencies in the WSA process

Participants in the WSA process are asked to gain a basic familiarity with the competencies assessed. In the course of their teaching and assessing during a year they are asked to give some thought to the extent to which, and the ways in which, the students are demonstrating the competencies. They might make notes during the course if they are so inclined, but such records are not required, and it is expected that many teachers can confidently make

informed judgements about some of the competencies for some students without recourse to formal reflection during the course or formal records. Towards the end of the course, teachers are asked to reflect formally and record their judgements in an Excel file.

The assumption behind the WSA software is that the judgements made in this assessment are stage related, in that they are made explicitly about students in specified grade levels at a specific point of time. The judgements are made on the basis of participating teachers' knowledge and experience of students at the specified level and they are based on what teachers know and expect of students at that level.

WS assessments are based on the fact that some students stand out, either because they exceed or do not meet usual and reasonable expectations. Students who do not meet basic expectations are judged to be below level one, and those who clearly distinguish themselves as high level performers are at level three. These suggestions imply that most students will fall in levels one and two on a particular competency because they meet basic expectations for the level but are not clearly distinguished. The distinction between Levels 1 and 2 is the difference between a basic level and some clear strength in a particular competency.

These presumptions can be expressed in another way to explain the process a teacher might use in making these competencies judgements. If a teacher is tempted to think a student is particularly strong or particularly weak in a competency, s/he is faced with a decision between Levels 3 and 2 and between Not Yet Level 1 and Level 1 respectively. If a student is not clearly distinguished then the teacher is deciding between Levels 1 and 2. In deciding between levels 1 and 2, teachers are looking for signs of strength that raise students above the basic expectation.
The WSA process is based on the three different roles of:

1. **Program Manager**;
2. **Assessing Teachers**; and
3. **Overall Assessors**.

The **Program Manager (PM)** organises the process and controls the software. Role 2 in the process is that of the **Assessing Teacher (AT)**. **Assessing Teachers** enter their assessments in to the Excel file containing the details of the students they are to assess that has been produced by the **PM**. Role 3 is that of the **Overall Assessor (OA)**. The **Overall Assessor** reviews the input of individual teachers through screen 4 of the software. As a result of this

review of the assessment level or levels suggested by the software for each competency, the **OA** either confirms or changes them to determine an overall result.

The WSA process for the Assessing Teacher

The role of the **Assessing Teachers** is to make level judgements on the competency performance of students. Assessing Teachers enter their assessments in columns G to P of the Excel spreadsheet.

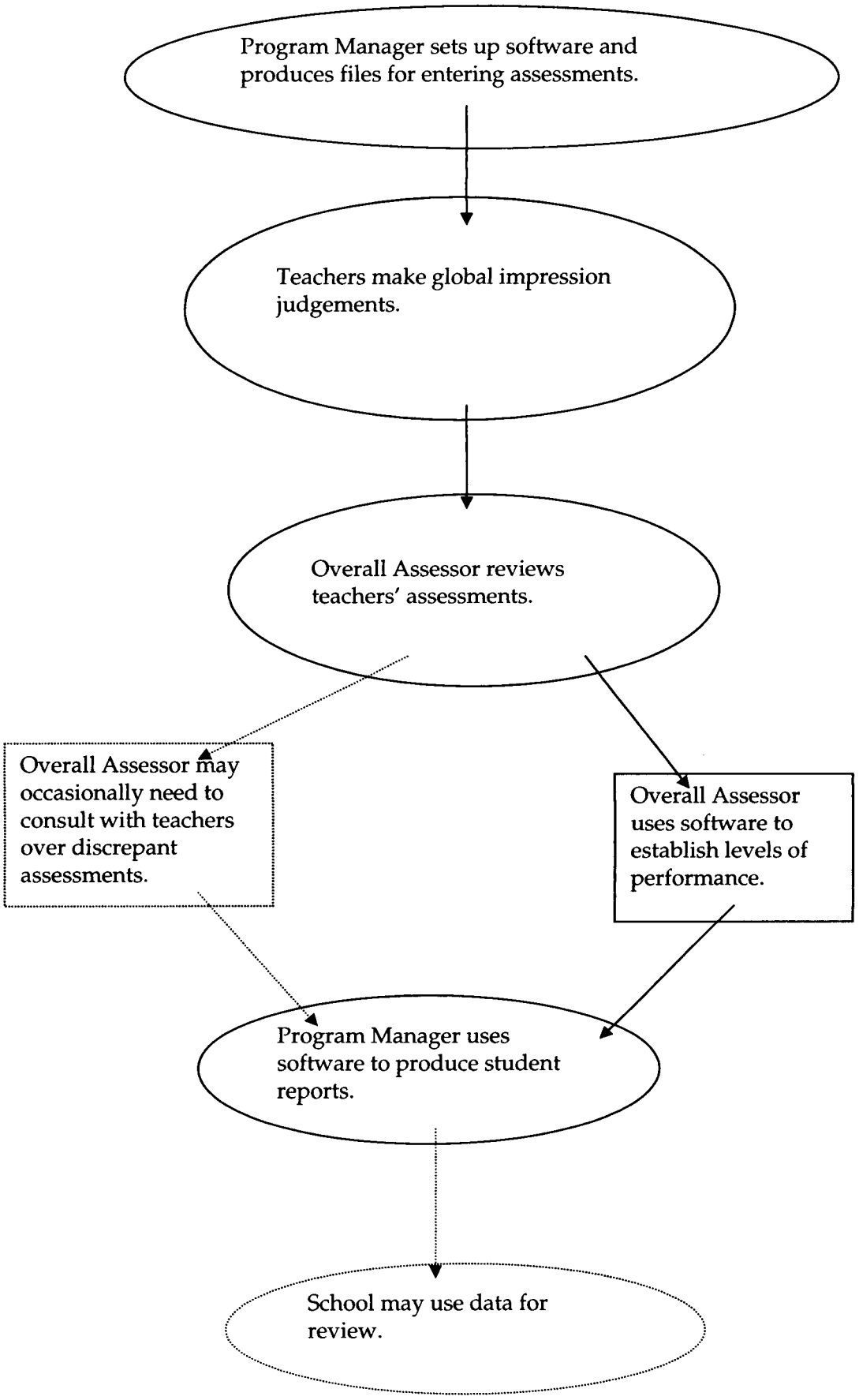
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	TeachID	Surname	FirstName	StudID	StdSurname	StdFirstName	KC1	KC2	KC3	KC4	KC5	KC6	KC7	KC8	KC9	KC10
2	1	Brown	Anne	1	Andrews	Angela	3	4	5	3	4	5	3	4	5	
3	1	Brown	Anne	2	Burke	Brenda	3	4	5	3	4	5	3	4	5	
4	1	Brown	Anne	3	Craig	Cathleen	3	4	5	3	4	5	3	4	5	

The WSA framework software allows the setting of different numbers of assessment levels. The default settings are shown in the table 39.

WSA Levels	**Level Code**	**Level Descriptors**			**#**
Level 3	**H+**	Very high	Very well developed	Exceptional	8
	H	High		Outstanding	7
Level 2	**M/H**	Medium to high	Well developed	Strong	6
	M	Medium		Some strength	5
Level 1	**B/M**	Basic to medium	Developing	Satisfactory	4
	B	Basic		Usually satisfactory	3
Not yet level 1	**NY/B**	Not yet to basic	Emerging	Not yet satisfactory	2
	NY	Not yet		Unsatisfactory	1

Table 39 The WSA default levels of performance codes and descriptors

Figure 28 An overview of the whole-school assessment process



The WSA process for the Overall Assessor

The role of the **Overall Assessor** is to produce a whole school report for each competency for a student based on the input from **Assessing Teachers**. After the assessments by teachers of students have been imported into the database, the results of the group of teachers assessing a student are automatically merged together so as to be accessible through screen 4 of the software (Figure 29) for the review of an **Overall Assessor**.

The **Overall Assessor** reviews the grade range proposed by the software and confirms or changes it. The **OA** can accept the suggestion of the software by clicking the **Accept Suggested Ranks** button. Or the **OA** may change the result by inserting results using the **Lower** and **Upper** dialogue boxes.

Screen 4 presents each of the competencies (up to 10 may be highlighted) as buttons at the top of the screen, and the teachers' assessments for each competency can be reviewed by clicking the appropriate button (C1 to C10). Below the name of the competency is the name of the student on the left and on the right is the level result (based on the input of different teachers) suggested by the software as the appropriate level or range of levels for the student on that competency.

The box at the bottom left corner of screen 4.1 displays the levels offered by teachers for the selected competency. The **OA** may review this information in making a decision to confirm or change the overall result suggested by the software. A level range suggested by the software appears to the right of the student's name. The result may range from one to three points on the 8 point WSA scale. Whether the suggested range is one, two or three points depends on the degree of agreement between the assessing teachers. If there is insufficient agreement between the teachers to give a meaningful report of three levels or less, the program will indicate this with the statement: **'A broad range of views'**.

Next to the heading **Result** in Figure 29 the software suggests a result for the student on the competency. In Figure 29 the suggested result is the range of Not Yet Basic to Basic. The **OA** can examine the assessments of individual teachers in the white box in the lower left corner of Figure 29. This box shows the names of the teachers who made the assessment in the first column, the Mean for all assessments by this teacher of this competency and the Grade for this student.

In Figure 29 Bella Grey gave a grade of NY/B and Anne Brown and Colin White gave assessments of B. The software suggests as outcome of NY/B to B. Notice that Bella Grey gave the lowest score for this students, and her mean score of two in all assessments of the competency is one point lower than that of Anne Brown and Colin White who have mean scores for this competency of 3.

The displaying of the mean for each teacher allows the **OA** to see easily whether a particular teacher has been harder or easier than others on a particular competency when reviewing the scores to make an overall assessment.

The **OA** may decide to accept the result suggested by the software by clicking the **Accept Suggested Rank** button. The **OA** may decide to change the suggestion of the software by choosing a lower and an upper result in the two dialogue boxes.

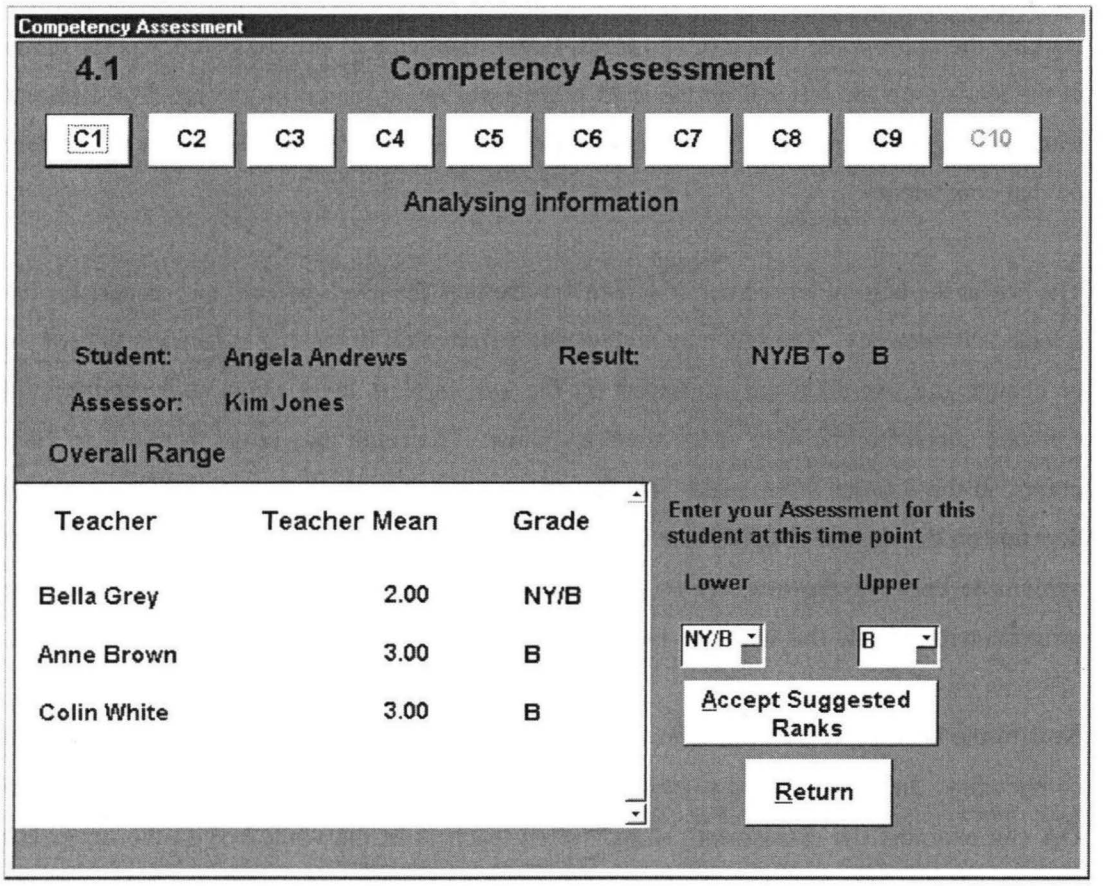


Figure 29 The Screen for the Overall Assessment

The Overall Assessor and dealing with discrepancies in the WSA process

All the individual teacher judgements made in the WSA process will eventually be reviewed by an **Overall Assessor** who will construct an overall school assessment on the basis of the judgements of the participating teachers. The program will offer the **Overall Assessor** a suggested overall assessment or an indication that agreement is insufficient and reconciliation is needed before a result can be suggested.

The decision-making process of the **Overall Assessor** will take into account the suggestions of the software, but the decision will not be an algorithm, like an average. The decision will take into account the perspectives of different teachers in respect to a particular competency and the pattern of all the other assessments of individual teachers.

The aim of the **Overall Assessor** will be to determine as many overall results as possible without requiring any further information or discussion. If the range of levels offered for an individual student by different initial assessments is within three points on the eight point scale, then the **Overall Assessor** can comfortably determine an overall result. If the range of initial assessments is more than three points on the scale, careful consideration will have to be given to the range of levels that might be appropriately offered, and a process of reconciliation of discrepant levels may have to take place before an overall result can be determined.

Reconciliation of discrepant groups of assessments will take place through consultation between individual teachers and the **Overall Assessor** in cases where a particular teacher seems to be at odds with other teachers. Some reconciliations may be conveniently managed as a result of such one to one discussion. Where there is significant diversity between all teachers involved, reconciliation may be attempted through group discussion.

A meeting of interested parties may be convened to consider difficult cases so as to see if a reconciliation can be made between the different perspectives. Teachers may also raise queries about particular overall assessments at such a meeting. If a consensus or a majority view cannot be reached about the appropriate level or range for a particular student, a result will not be recorded for that student.

If a level of performance is to be reported and the software says that there is a broad range of views among the assessing teachers, it is the role of the **Overall Assessor** to make a decision

about what might be reported, or to consult with **Assessing Teachers** to see if a consensus can be reached.

In attempting to make a decision the **Overall Assessor** will look to see if any particular teachers are at odds with the others, and if this seems to be the case the **Overall Assessor** may:

- assess the perspective on the competency given by the teacher's subject area; and
- assess the mean level of the assessments given by any particular teacher in comparison with other teachers to see if they are consistently higher or lower than other teachers.

On the basis of this information the **Overall Assessor** may decide to discount the assessment level of one or more teachers in the light of the agreement between other teachers. On the other hand the **Overall Assessor** may choose to discuss the particular student with teachers whose assessments are inconsistent with those of other teachers to see if the inconsistent teachers are prepared to reconsider their view in the light of those of their colleagues. In some cases the **Overall Assessor** may determine that an individual needs to be discussed by all assessing teachers if an Overall Level Judgement is to be reported.

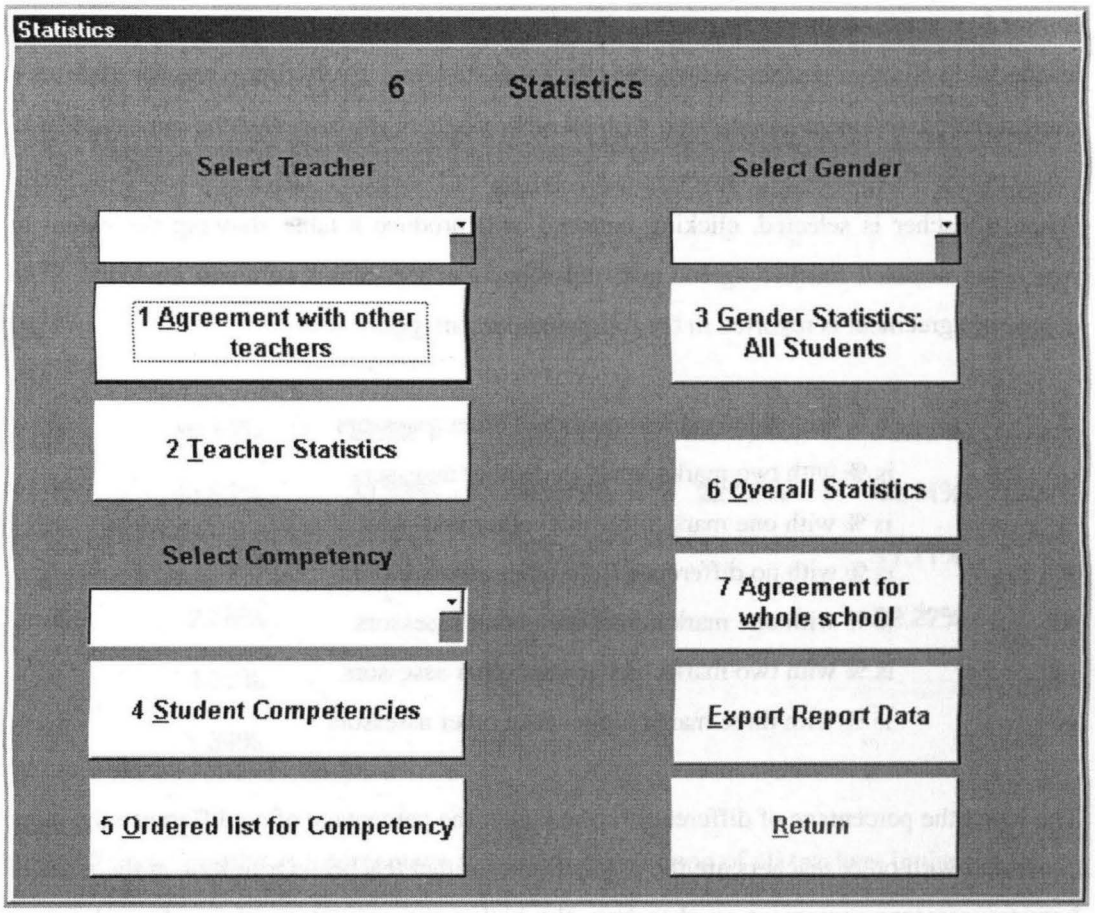
The **Overall Assessor** may change a result suggested by the software by inserting the preferred range in the two overall range boxes in the bottom left corner of screen 4.1. Such insertions here will override the suggestions of the software.

Student reports

After the work of the **Overall Assessor** is completed, two kinds of report can be produced. The **Levels** option gives the level or levels determined by the **Overall Assessor** only. **The Levels and Range** option produces both the level determined by the **Overall Assessor** and shows the mean and standard deviation for all students in the cohort as a comparison with the result of the particular student. (See the Level Mean row below each assessment on the student report example in Figure 27.) This row shows the mean score for all students at this level at this time as a black bar in the middle and one standard deviation (approximately 33% of the student scores) from the mean as two thinner black bars on either side of the mean.

Analysing the data

The Statistics facility of the WSA software can produce summary statistics based on the level judgements of the assessing teachers. The program analyses the level judgements of individual teachers and the school as a whole to produce the following statistical output.



The results given in the following examples of the software output are fictional.

Agreement With Other Teachers

Jim Black

Grade Difference	Competencies 1	2	3	4	5
-2	15.38%	7.69%		30.77%	
-1	23.08%	30.77%	33.33%	23.08%	
0	53.85%	23.08%	33.33%	38.46%	46.15%
1	7.69%	15.38%	33.33%	7.69%	23.08%
2		15.38%			23.08%
3		7.69%			7.69%

Button 1 on screen 6 shows the results as a percentage of a pairwise comparison of a selected teacher with all other teachers who assess the same students. Every time a teacher assesses a student their assessment is compared with all other teachers who assessed the same student.

When a teacher is selected, clicking button 1 will produce a table showing the extent to which the selected teacher agreed with all other teachers about common students. The degree of agreement is reported in the following percentages.

-3	is % with three marks lower than other assessors
-2	is % with two marks lower than other assessors
-1	is % with one mark lower than other assessors
0	is % with no difference from other assessors
+1	is % with one mark higher than other assessors
+2	is % with two marks higher than other assessors
+3	is % with three marks higher than other assessors

The lower the percentage of difference or the higher the percentage of no difference the more consistent with other teachers are the assessments for that teacher. Or to look at the issue of between assessor agreement another way, the higher the percentage of 0, +1 or -1 for an individual teacher the better.

Whole school agreement statistics

Agreement Between Teachers For Whole School

	Competencies				
Grade Difference	1	2	3	4	5
0	30.56%	22.22%	36.11%	77.78%	22.22%
1	45.83%	33.33%	33.33%	22.22%	44.44%
2	18.06%	44.44%	23.61%		11.11%
3	2.78%		6.94%		22.22%
4	1.39%				
5	1.39%				

Button 7 shows results as a percentage a pairwise comparison of all teachers in the school.

0	is % with no difference between the pairs
1	is % with one point difference between the pairs
2	is % with two points difference between the pairs
3	is % with three points difference between the pairs

The lower the percentage of difference or the higher the percentage of no difference the more consistent and meaningful are the assessments produced by the school as a whole.

Gender statistics

Button 3 gives the average score and standard deviation for male and female students after a gender has been chosen from the dialogue box.

Ordered list for competency

Button 4 gives the average score for each student in a list for each competency from highest to lowest when a competency has been selected from the **Select Competency** box.

Student list for Collecting and analysing

Student Name	Mean
Bolton , Bernie	5.75
Jastrow , Jennifer	5.50
Cook , Christine	5.00
Geoffrey , Graham	5.00
Kristen , Kate	5.00
Landsdown , Leo	5.00
Morris , Meg	5.00
Dodson , Darryl	4.83

Teacher statistics

Button 2 gives the mean and the standard deviation for a selected teacher. One would not expect to see large differences in mean and standard deviation between different teachers.

Statistics For Jim Black

Competency	Mean	Std Deviation
Collecting and analysing	4.33	1.15
Communicating	5.00	1.73
Planning and organising	5.00	
Working with others	4.33	1.15
Using mathematics	5.67	0.58
Using technology	5.00	
Solving Problems	5.00	

Overall statistics

Statistics For Whole School

Competency	Mean	Std Deviation
Collecting and analysing	4.86	0.93
Communicating	4.93	0.87
Planning and organising	4.80	0.84
Working with others	5.07	0.81
Using mathematics	4.95	0.94
Using technology	5.60	0.55
Solving Problems	5.40	0.55

The **Overall Statistics** button 6 gives the mean and the standard for all teachers of all subjects. The means for different competencies will give the school a summary of the comparative strengths and weaknesses of their students. The Standard deviation will give an indication of the dispersal of the scores.

The statistics above can be used by the school to monitor the performance of its students and the assessments made by individual and groups of teachers. The following questions indicate some of the issues this analysis might be used to elucidate.

- In which competencies do the students in our school perform strongest and weakest?
- Which individual students perform strongest and weakest in different competencies? Which students are excelling and which are at risk?

- How might the school's programs be developed to deal with any comparative weakness in student performance in particular competencies?
- To what extent do our teachers agree in the way they see the students?
- Are there some things that they agree about and some that they differ about? Where might they develop greater agreement?
- Why might there be more or less agreement between different teachers?
- Might individual teachers change their perspective to better align themselves with their colleagues?

As can be seen from the outline above, the WSA Software is designed around and supports an assessment process. The software brings the judgements of different teachers together so that a single integrated report can be produced for a student. The software also allows convenient analysis of the data from individual teachers and the school as a whole.

11
An Overview and Conclusions

The Key Competencies in Operation

As we have seen in the preceding chapter, using the Key Competencies effectively must entail circumscribing their breadth and sharpening their focus. Furthermore, this circumscription and focusing should be done so as to increase the differentiation and minimise the overlap between the Key Competencies (in the manner of the kinds of differential ability theory discussed in Chapter 7).

The SBKCLAP trial showed that significant numbers of participants could not assess KC3 Mathematics, KC6 Technology or KC4 Cultural understanding, and it seems that these matters are fairly specific rather than cross-curricular. Breaking communication into writing and speaking in the trial was successful, and it shows that an efficient cross-curricular assessment can be produced for these basic generic skills using the methods developed for the trial. This is a very significant result in itself, but this assessment method can deal with more than the generic skills of oracy and literacy.

Interpreting Problem solving, Planning and organising and Teamwork as matters of general attitude or stance was successful in this trial. In Planning and organising and Teamwork, assessing teachers made judgements of kinds of activity or attitudes. Such judgements are not usual in academic assessments. Where participants may well have been uncomfortable about making individual judgements about the personal skills or attitudes of students, they were quite comfortable contributing to a group judgement on such matters, or at least they were happy doing so in a trial situation.

Interpreting Problem solving as an attitude or stance was successful in that it circumscribed a very vague term, and it moved participants towards judgements about the personal qualities of students in a measured or qualified fashion. Making a judgement about Problem solving as an attitude had the benefit of seeming to deal with a kind of activity while actually edging towards an attitude or personality trait. It is this guarded movement towards the assessment of attitudes and personality traits that turns out to be the distinctive, pragmatic strength of the Mayer conceptualisation.

Theory, Research and the Key Competencies

It has been noted at various points in this discussion that the Mayer Key Competencies were atheoretically conceptualised. It might be argued that there is no reason why the Mayer Committee should have concerned itself with either empirically grounding its work in or being concerned to produce an ability theory that would meet psychometric standards. The focus of the Mayer Committee was pragmatic. They were concerned with what was needed for coping with the future, and they wanted to produce a set of conceptions that would be a worthwhile and practical stimulus to a re-orientation of education and training. The Mayer Committee wanted conceptions that would make sense to educators, employers and the community in general. Unlike the other work-related conceptions discussed in Chapter 2, the Mayer Committee also wanted conceptions that could be assessed.

Given the powerful interests involved, few were prepared to directly challenge the avowed purpose of the Mayer Committee during its work and after it had finished. Instead there were more indirect challenges, on the basis of 'theory and research', to the soundness of the Mayer proposals. The Mayer Committee offered no more in response to these challenges than to assert that its competencies were to be generic and transferable and that they included foundation knowledge. These mere assertions were not an adequate response to the challenges made by critics to the Key Competencies.

It has been argued here that the challenges made to the Mayer Committee proposals on the basis of theory and research are not decisive and are in a sense irrelevant to the fundamental issues. As well, it has been argued here that there are general capacities that are used in all kinds of activity, and that these general capacities are particularly important in learning new things. It has also been argued here that these general capacities are more effectively and appropriately discerned in non-specialist activities that aim to assess the aptitude to learn rather than the particulars of what has been learned in the past. These general capacities or abilities are most effectively and appropriately viewed as aptitudes rather than achievements or expertise. Hence it is also argued here that the notion of generic or key competencies is confused and confusing because competencies can only be understood as a matter of domain-specific knowledge and expertise.

The Mayer Committee made none of these distinctions, and it blandly ignored challenges to its Key Competencies. It might have been thought by the Mayer Committee that they did not need to deal with such theoretical arguments because they were practical people

pursuing practical ends. However, the indifference of the Mayer Committee to the criticisms of the theorists meant that the stance of the Committee remained confused and confusing, and that the work of the Committee was less practical than it could and should have been because of that confusion. The history of the Mayer Key Competencies shows that good theory is practical. The work of the Mayer Committee would have been significantly more practical if it had been more theoretically coherent and comprehensive.

The Mayer Committee should have more carefully and rigorously worked out what kinds of constructs they were concerned with, and they should not have confused matters by loosely using the fashionable terms 'competence' and 'competencies'. Similarly, they should not have made an explicit declaration that they were concerned with knowledge and skills and rejected the assessment of attitudes and values while at the same time blandly presenting for assessment the non-academic and attitudinal constructs of Planning and organising and Teamwork. The Mayer Committee should have offered a more rigorous examination and explanation of what it proposed. Had it done so, it would have produced a stronger and more practical set of proposals.

Key Competencies and Personal Skills

There is, however, one respect in which the work of the Mayer Committee was practical and in which its work has proved to be robust. As shown in the SBKCLAP trial, the Mayer Key Competencies can be conceptualised coherently and implemented effectively. Leaving aside the reservations mentioned above, the Key Competencies of Planning and organising and Teamwork (and the interpretation of Problem solving as an attitude) turn out to be very well framed by the Mayer Committee for implementation in an assessment. These notions turn out to be useful and significant expansions of the scope of educational assessments without excessively challenging the confidence of teachers or the legitimate scope of public assessment in a pluralist community.

It has been argued here that Planning and organising and Teamwork are not usefully or realistically thought of as particular sets of knowledge and skill. They are not domains, and they do not offer themselves as a curriculum or a course of study. Contrary to what the Mayer Committee claimed, these constructs are personal skills that contain a significant element of the attitudinal in them. These more attitudinal and less cognitive constructs identified by the Mayer Committee are phrased as kinds of activity, and hence touch on matters of attitude and personality in a careful and guarded fashion.

In comparison with the other conceptions of work-related skills discussed in Chapter 2, the Mayer Key Competencies are the most practical and realistic. The more cognitive Key Competencies should be tightened and rationalised, but what of the more attitudinal and less cognitive? What needs to be, or might be added, to the Mayer Key Competencies in these respects? As it turns, out this question has been raised as part of a new wave of interest in Australia in notions of work-related skills in the early years of the first decade of a new century.

Further Developments in Generic, Work-related Skills

The Kearns Review of Generic Skills

The Australian National Training Board, through the National Centre for Vocational Education Research, commissioned Kearns in 2000 to consider the state of play on the matter of generic skills. In a review of research entitled *Generic Skills in the New Economy*, Kearns examined the various work-related skills projects discussed in Chapter 2 (Kearns 2001). He emphasised the comparative narrowness of the Mayer conception, as outlined in Chapter 2, particularly in terms of it exclusion of attitudes and values. Kearns attempted to make a comprehensive map of the possibilities for defining generic skills. He represented his conclusions in Figure 31.

The relationship between the elements of Kearns' cluster of generic skills is hard to comprehend and interpret, and the distinctions the model offers are unclear. In particular Kearns' model does not distinguish skills from attitudes or cognition from conation and affect.

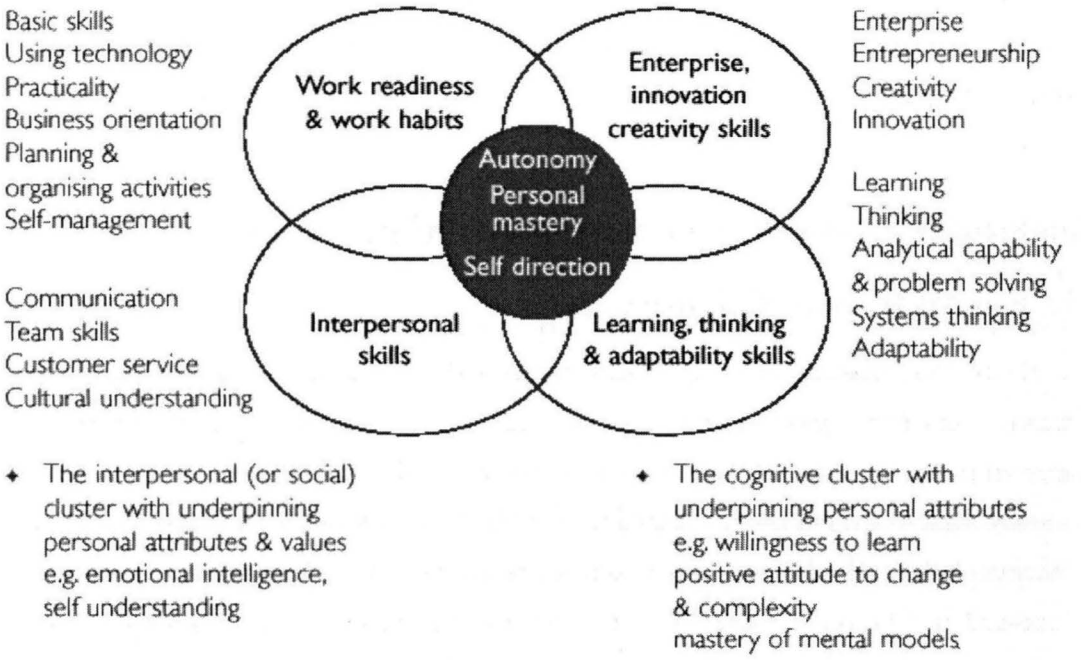


Figure 30 Kearn's Clusters of Key Generic Skills

Employability Skills

In 2001, the Department of Education Science and Training and the Australian National Training Authority funded a project to analyse and report on current business requirements for 'employability skills'. This project, conducted by the Australian Chamber of Commerce and Industry and the Business Council of Australia with the (ACCI/BCA), was completed in March 2002 (ACCI and BCA 2002).

The employability skills project involved a research paper by Curtis entitled 'Employability skills for Australian industry: literature review and framework development' (Curtis 2001). Like Kearns, Curtis recognised that the Key Competencies were more circumscribed than other conceptions of work-related skills. Curtis developed a classification or framework (Table 44) of the kind presented in this discussion for analysing the possibilities into the categories of basic skills, intellectual abilities and personal attributes.

Table 40 Composite Skills Set Identified from National Frameworks of Employability Skills and Employers' Views by Curtis

Basic Skills	Intellectual Abilities	Personal Attributes
Foundation skills	**Thinking skills**	**Continuous learning**
Listens and understands and speaks clearly and directly	able to make decisions	acknowledges the need to learn in order to accommodate change
Understands written documents and writes clearly	capable problem-solver	open to new ideas and techniques
Understands tables of figures, interprets graphs, able to calculate	innovative – adapts to new situations	is prepared to invest time and effort in learning new skills
	creative	
Information and communications technology	**Systems understanding**	**Personal attributes**
is willing to use a range of technologies	knows own role in the work situation	has positive self esteem
can use technology to seek, process, and present information	understands interrelationships among workplace processes and systems	understands that own actions influence others
	can diagnose system (process) deficiencies*	is self-manager, resourceful, shows initiative and effort
	can design, implement, and monitor corrective actions*	displays sense of ethics including integrity and honesty
		accepts responsibility for own actions
		seeks and accepts feedback
	Project management	**Interpersonal skills**
	is able to manage own time and to seek needed resources to complete set tasks	shows cultural sensitivity
	sets goals and engages others in achieving those goals	committed to client service
	establishes clear project goals and deliverables*	able to negotiate
	allocates people and other resources (eg budgets, materials, space) to tasks*	works well with others, individually and in teams
	sets time lines and coordinates sub-tasks*	shows leadership*
	is able to adapt resource allocations to cope with contingencies*	can develop a strategic vision, set goals, and monitor performance*
		communicates goals and targets, engages and enthuses subordinates towards a shared vision*

Note: * Indicates that the attribute is expected of experienced workers, but not of new entrants to the workforce Curtis (p.51).

Other aspects of the Employability Skills project involved consultation with business. The report of this consultation, *Employability Skills for the Future* (ACCI and BCA, 2002), said that business and industry required a broader range of skills than the Mayer Key Competencies. The research found that business placed particular value on certain personal attributes that were not part of the Mayer Key Competencies. Business indicated that personal attributes were as important as other employability skills, and that learning skills and self-management skills should also be included in the new framework.

Table 41 The Employability Skills Framework

Employability skills Framework	
Personal attributes	
• loyalty • commitment • honesty and integrity • enthusiasm • reliability • personal presentation • commonsense • positive self-esteem • sense of humour • balanced attitude to work and home life • ability to deal with pressure • motivation • adaptability.	
Key Skills	
Communication	skills that contribute to productive and harmonious relations between employees and customers
Team work	skills that contribute to productive working relationships and outcomes
Problem-solving	skills that contribute to productive outcomes
Initiative and enterprise	skills that contribute to innovative outcomes
Planning and organising	skills that contribute to long-term and short-term strategic planning
Self-management	skills that contribute to employee satisfaction and growth
Learning	skills that contribute to ongoing improvement and expansion in employee and company operations and outcomes
Technology	skills that contribute to effective execution of tasks.

The Kearns model does not make distinctions clear, and one might question the distinction between basic skills and intellectual abilities (what could be more basic than intellectual abilities?) made by Curtis. These researchers do make the more cognitive/less cognitive and more attitudinal/less attitudinal distinctions made in this discussion. Although Curtis produced quite a clear framework, because it is not focused and circumscribed the prospect of actually implementing it is daunting.

Both Kearns and Curtis note the circumscribed nature of the Mayer Key Competencies, and they offer a much broader frameworks in line with DeSeCo and the other work-related frameworks reviewed in Chapter 2.

The personal attributes of the employability skills framework are a very specific wish list of personal traits that employers would want in employees. There seems to be no consciousness in the ACCI/BCA project that such personal attributes are difficult to deal with in educational programs and in any kind of assessment. However the Key Skills of the Employability Framework reflect the Key Competencies. All of the Mayer constructs (other than Using mathematical understandings) are included in the employability skills with the addition of initiative and enterprise, self-management and learning skills. Although the employability framework claims that initiative and enterprise and self-management are skills, they are more appropriately seen as attitudinal constructs. The notion of 'learning skills' is dauntingly broad.

In the light of the analysis offered and the experience described in this discussion, it can be seen that the Mayer Competencies turn out to be carefully framed and circumscribed in some important senses, and that they offer a more realistic basis for an assessment than any of the other models of work-related skills that are reviewed here. The recent attempts to conceptualise work-related skills, and the SBKCLAP experience, leads to the conclusion that the Mayer Key Competencies are judiciously specific and circumscribed.

Towards a Taxonomy of Abilities

In the light of this discussion, how might one envisage a framework or taxonomy of abilities and skills?

The Foundation Skill of Language

The notion that literacy and oracy are generic or basic skills is fundamental to the DomPro model sketched in Chapter 7. Literacy and oracy are based on knowledge and skills that can be used in all kinds of activity and all kinds of domain. Numeracy, on the other hand, is based on knowledge and skills that can be applied across different domains, but it is arguably a knowledge domain in itself. Literacy is a generic skill whereas numeracy can be seen as a more of less specific body of knowledge. They are not equivalents.

In seeing literacy as a generic skill we see that literacy is:

- a kind of activity;
- based on a body of knowledge;
- that can be applied to all kinds of domains.

Speech operates at the level of our identity as human beings, and reading and writing are almost as fundamental as speech. In these terms it is difficult to see other kinds of task or kinds of knowledge that are analogous to oracy and literacy. In comparison with literacy and oracy, numeracy is a knowledge domain. We might think of reasoning as a general and transdomainal activity, but there is cause for asking whether reasoning is more or less the same across domains or whether it differs in different kinds of domains or whether there are different kinds of reasoning. While one might accept the notion that reasoning is a core skill, there are reasons for not accepting (as in Cattell's model of cognition) abstract and formal reasoning as the central, transdomainal ability.

Differentiating or nuancing reasoning is the purpose of the DomPro model. In its fluid intelligence type tasks, the psychometric tradition has focussed on a kind of abstract and decontextualised reasoning, but it has not usefully theorised other kinds of reasoning. Verbal reasoning is the other major psychometric construct that is distinguished from abstract reasoning and spatial/visual reasoning. Verbal reasoning is the first of the abilities discussed by Carroll and a linguistic intelligence is one of the basic cognitive elements of Gardner's scheme, but Carroll's verbal reasoning (like Gardner's linguistic intelligence) is a is not a subtle or nuanced concept.

The Psychometric Notion of Verbal Reasoning

Carroll (Carroll 1993) offers verbal factors as a kind of 'baseline for the consideration of other factors' (p.145):

> It is often noted that the most frequently found factor in factor-analytic studies is a 'Verbal' (V) factor (p.151).

Although there have been significant attempts to differentiate verbal ability, Carroll concludes that knowledge of vocabulary and reading comprehension are the fundamental aspects of psychometric verbal ability. Carroll also notes that since this verbal ability is commonly a kind of general knowledge, it is the major construct of crystallised intelligence against which fluid intelligence is distinguished.

It has not been uncommon in the psychometric tradition to equate language development with cognitive development in general, but Carroll does not support such a specific equation:

> It has been apparent that the development of language skills is substantially related to the development of more general cognitive skills, in that measures of language skills tend to be substantially correlated with measures of other cognitive skills, and various primary factors are dominated by more general factors. It has as often been noted (e.g., Terman, 1916) that measures of vocabulary are among the best predictors of general intelligence. It is tempting to speculate that general intelligence is identical to rate and extent of language development, but such a speculation is only weakly supported in research. Language development as customarily measured (e.g., by vocabulary tests) is only one aspect of cognitive development, and its correlations with intelligence measures are not so high as to suggest that language development is the same as general cognitive development. There are many influences that govern individuals' rate and extent of language development - exposure to increasing levels of language complexity through exposure to model speakers, reading of increasingly difficult material, etc. Special traits associated with rates of learning to read may make estimates of general intelligence obtained from printed tests inaccurate. On the other hand, it is probable that the level of general cognitive development that is or can be attained by an individual at a given age tends to set limits on the level of language development that can be attained at that age (p193).

It is apparent that the notion of verbal ability presented by Carroll is very broad and undifferentiated. It is a matter of how much language an individual knows and how accurately an individual comprehends and interprets language.

Knowledge of word meanings is a comparatively simple matter in assessment terms. It can be directly tested, and there is a fairly clear hierarchy of knowledge related to the frequency of word use. Understanding words as concepts and in conceptual relationships with each other is more complex, and it moves from recall of knowledge into reasoning with or about that knowledge. Reading is a complex relationship of word knowledge, conceptual understanding and understanding of cultural context to make meaning. In its full complexity, reading is a kind of reasoning, but it is a largely tacit kind of reasoning, and the different aspects of reading may add up to different kinds of reasoning. Clearly the kind of reasoning involved in reading is significantly different from the abstract reasoning of fluid intelligence tests like Raven's Matrices.

Kinds of Verbal Reasoning

Although the term verbal reasoning is commonly used in the psychometric tradition, it covers quite different kinds of tests. In the Scholastic Achievement Test (SAT) of the Educational Testing Service, for instance, verbal reasoning means vocabulary knowledge, verbal analogies and reading comprehension. There is little or no reasoning in the SAT sub-

test of vocabulary knowledge, and the verbal analogies and the reading comprehension seem to be quite different kinds of reasoning.

The kind of verbal analogies that are a standard part of tests like the SAT are the equivalent in language of other non-verbal, fluid intelligence tasks. They offer candidates the pair of words '*doctor : patient*' followed by '*lawyer : ?*', and ask candidates to choose a word to pair with '*lawyer*' that matches the relationship of doctor and patient. What is striking about these exercises is that they are not the processes of analogising that are fundamental to thinking. They are not a process of finding a basis for a relationship between ideas or things that illustrate characteristics. (Is the relationship between a doctor and a patient like that between a parent and a child?) Verbal analogy test items are based on arbitrary comparisons, and one can see this arbitrariness by simply changing the pairs. For instance, the whole issue changes when the task becomes doctor is to patient as taxi driver is to passenger. The match can still be made, but the grounds for the match have been changed. On the other hand, actual analogising would be to ask which is the better comparison for the relationship between a doctor and a patient: that between a lawyer and a client or that between a taxi driver and a passenger? The example of verbal analogy items shows the weakness of the notions of verbal reasoning in the psychometric tradition and the failure of that tradition to adequately differentiate kinds of reasoning.

The Personal Skills of Planning and Organising and Teamwork

On the other hand, there are notions like Planning and organising and Teamwork that are quite unlike literacy or verbal ability. While Planning and organising and Teamwork are kinds of activity, they can actually cover such a diverse range of things that it is difficult to see them as more or less coherent kinds of activity. The standard argument we have seen about such general notions is that they have no meaning in general and have to be contextualised. Such an argument claims that there is no such thing as planning and organising or teamwork in general: that what is meant by teamwork depends on what kind of team is involved and what kind of work is at issue. Similarly, what is meant by planning and organising depends on the kind of planning and the kind of organisation at issue. Planning and organising a military campaign is not the same as planning and organising a research project. Teamwork on a battlefield is not the same as teamwork in a research institute. The particulars of these different activities are, of course, crucial to competent performance. (A researcher who might be very good at planning research projects might be overwhelmed by the challenge of planning a battle.) These differences and distinctions are not in dispute.

Planning and organising and Teamwork differ from literacy in that they are not more or less specific bodies of knowledge and skill. A specific program can be developed to teach how to read and write, but a specific program in Teamwork or in Planning and organising is not realistic (and very rarely attempted) because Teamwork and Planning and organising are not more or less specific bodies of knowledge or skills. It is not appropriate to think of Planning and organizing and Teamwork as bodies of tacit knowledge and skills like reading and writing, let alone thinking about them as though they are a body of knowledge and skills like mathematics.

This difference between a body of knowledge and skills and general notions like Planning and organizing and Teamwork prompts theorists like Weinert to the conclusion that such general notions are empty and meaningless. Similarly such considerations prompt Bowden and Masters to the view that generic capacities are only meaningfully elicited through discipline-based capacities.

One can readily agree that, as generalisations of kinds of tasks, or as sets of skills, Planning and organising and Teamwork have very little meaning. Such general notions however, do have meaning as abilities, that is as generalisations about kinds of capacity that can be deployed in different kinds of activity. A theory of abilities is a notion of different kinds of capacity that are brought to bear on different kinds of task, and a theory of abilities aspires to, and should have, general meaning and applicability. Such a theory of abilities aims to define general capacities that underpin learning and performance, and that are not limited to particular kinds of tasks and are not matters of particular kinds of knowledge. It has been argued here that these general abilities are best understood as aptitudes to learn.

As shown in chapter 5, the notion of a class of tasks definition of abilities and competencies is attractive because it seems to be a matter of observable performances, but the class of tasks definition of abilities and competencies is very difficult to construct, particularly for more attitudinal kinds of performance.

- What are these general kinds of attitudinal tasks?
- What makes them a class?
- How do we know that they are a class?

Although Weinert concludes that the class of tasks approach is the functional one, we have noted that he does not deal with any of these issues.

Planning and organizing and Teamwork are not readily thought of as classes of tasks or bodies of knowledge and skills. Thus when one is looking for generalised abilities, the most

useful way of envisaging them is as kinds of cognitive capacity. The argument presented here is that abilities can be seen as generalised if they are not tied to particular tasks or knowledge and skills. Hence, the meaningful way to envisage generic competencies is as kinds of ability rather than as kinds of tasks.

What conclusions might we draw from this discussion about generic ability and skill constructs?

Towards a Comprehensive Theory

Table 42 attempts to map the range and kinds of issues that may be considered in conceptualising abilities, skills and attributes. It offers a classification of generic skills, domains, domain-related skills, attitudes, qualities and values.

Table 42 A Classification of Generic, Domain-related and Personal skills

Generic skills	Domains and related skills	Personal qualities, attitudes and values
Generic enabling skills Writing skills Oral skills Reading skills (Analysing and interpreting skills) Numeracy	**Domains** Mathematics Science Technology Social science Humanities Arts	**Personal qualities** Leadership Initiative Flexibility and adaptability Reflection and evaluation Independence Responsibility
General reasoning skills Logical reasoning Critical reasoning Informal and plausible reasoning	**Domain related reasoning skills** Reasoning about scientific and technological issues Scientific and technological understandings Reasoning about cultural issues Cultural understandings Reasoning about personal and interpersonal issues	**Attitudes** Planning and organising activities Teamwork Problem solving approach Values Trust worthiness Responsibility
More cognitive		**More attitudinal**
Testable as an ability	**Performance assessment**	**Not testable as an ability**

The classification of constructs in Table 42 is differentiated in terms of being more cognitive at one end and less cognitive and more attitudinal at the other end.

The personal qualities, attitudes and values of the third column of this table are the matters that are added to the DomProC figure presented in Chapter 8 to amount to the following DomProAt figure.

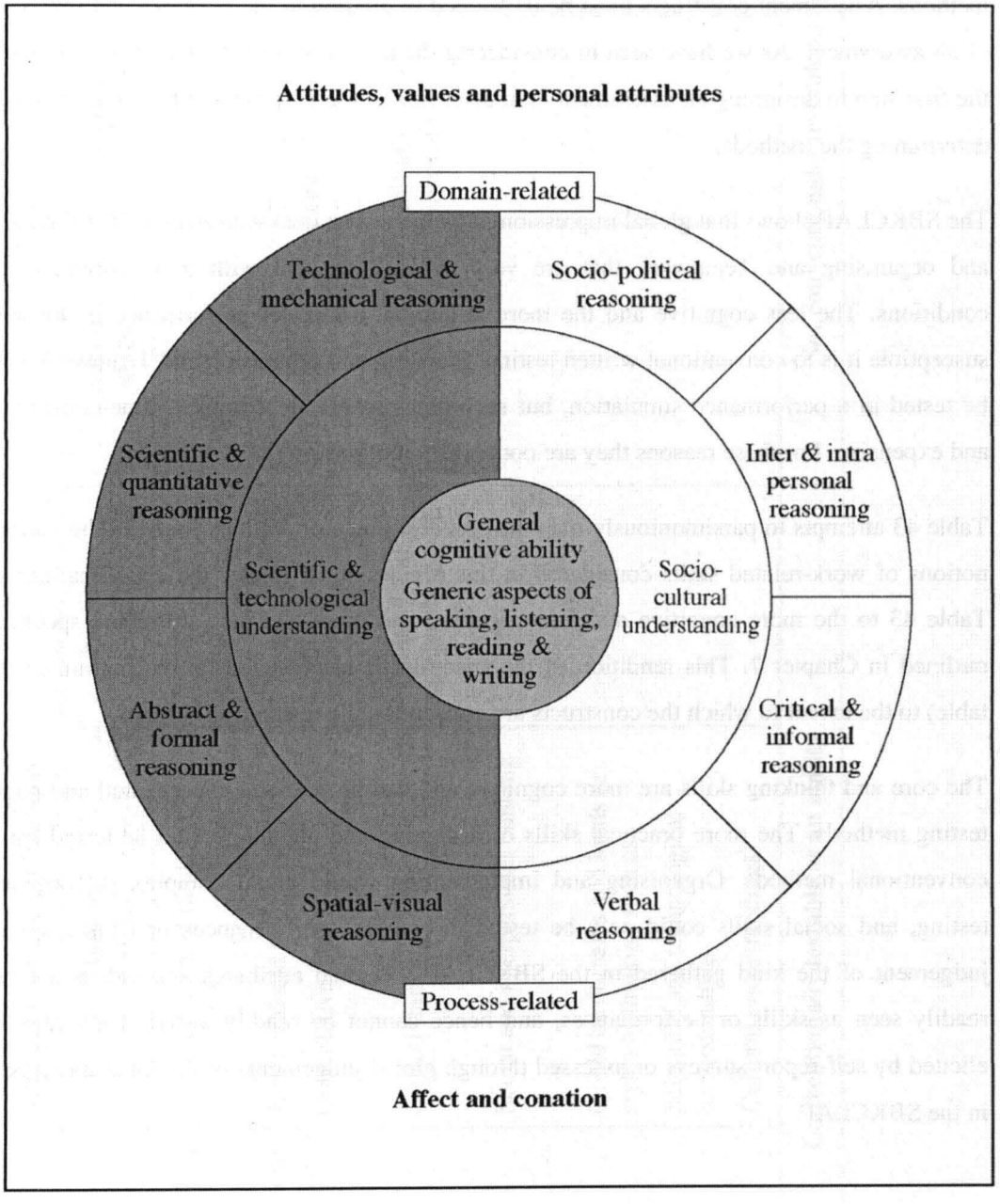


Figure 31 The Domain and Process-related Differentiation of Cognitive Ability Nested in Conation and Affect (The DomProAt model)

What do Table 42 and Figure 32 suggest about assessment possibilities?

How might these different skills and qualities be assessed?

The Assessment of Generic Skills

When one considers the DomProAt as a basis of assessment, one is reminded of the way in which assessment constructs should be conditioned and constrained by possible assessment methods. Assessment constructs must be developed in relation to the purposes and functions of an assessment. As we have seen in considering the assessment of the Key Competencies, the first step in designing an assessment is in determining the purpose and function, and then determining the methods.

The SBKCLAP shows that global impression judgements can deal with notions like Planning and organising and Teamwork that are very difficult to deal with in constrained test conditions. The less cognitive and the more attitudinal a kind of performance is, the less susceptible it is to conventional written testing. Planning and organising and Teamwork may be tested in a performance simulation, but such assessments are complex, time-consuming and expensive. For these reasons they are not realistic options.

Table 43 attempts to parsimoniously map the broadest range of abilities found in the various notions of work-related skills considered in this discussion. It relates the classifications of Table 43 to the more cognitive and less cognitive and more and less attitudinal spectrum outlined in Chapter 7. This rendition of the spectrum is also related (at the bottom of the table) to the extent to which the constructs are conventionally testable as abilities.

The core and thinking skills are more cognitive and can be assessed using pencil and paper testing methods. The more practical skills of designing and planning could be tested using conventional methods. Organising and implementing would entail complex performance testing, and social skills could only be tested in complex performances or from a global judgement of the kind gathered in the SBKCLAP. Personal attributes and values are not readily seen as skills or performances, and hence cannot be readily tested. They may be elicited by self-report surveys or assessed through global judgements of the kind undertaken in the SBKCLAP.

Table 43 A Spectrum of Kinds of Skills, Attributes and Values

More cognitive / less attitudinal			Less cognitive / more attitudinal		
Core skills	**Thinking skills**	**Practical skills**	**Social skills**	**Personal attributes**	**Attitudes and values**
(generic and enabling)	(kinds or styles of thinking)				
Speaking, reading and writing	Analytical and logical thinking Quantitative thinking	Designing and planning	Team work and collaboration	Flexibility and adaptability	Trust worthiness
Numeracy	Critical thinking	Organising and implementing	Leadership and independence	Initiative and innovation	Responsibility
	Social and interpersonal thinking				
	(could be domains and kinds or styles of thinking such as Scientific and technological understandings Cultural understandings)				
Conventionally testable as an ability			**Not conventionally testable as an ability**		

The SBKCLAP demonstrated that practical and social skills and personal attributes can be assessed by global impression judgements even though such practical and social skills and personal attributes are not readily testable.

On the other hand, what might we conclude about the more cognitive and conventionally testable end of the spectrum?

We have seen consistently in this discussion that abilities interact with different kinds of content and knowledge. We have also seen that abilities offer the basis for broadest generic constructs. Might different kinds of ability be integrated with different kinds of topics or themes? And what are the conventionally assessable general cognitive skills?

For instance, literacy is eminently testable, but the testing of literacy can be problematic in that literacy can be conflated with other kinds of cognitive ability. Ideally one would wish to separate other cognitive abilities from literacy and further, one would wish to differentiate kinds of cognitive ability.

While one might hope for assistance from the psychometric tradition with this differentiation, this body of research has not been particularly subtle in its definition of kinds of thinking. As we have seen above, the psychometric tradition offers the notion of abstract and formal reasoning. However the notion of verbal reasoning in the psychometric tradition is somewhat problematic in that it too readily conflates knowledge of language with reasoning. The other major construct offered by the psychometric tradition is that of spatial/visual reasoning, but this is a fairly specialised kind of thinking, or a kind of perception.

Taking abstract and formal reasoning as a notion, what can one contrast it with? One answer to this question that is offered by the psychometric tradition is the notion of social cognition: that is, the notion of thinking about socio-cultural issues. Rather than using this more generalised notion of social cognition, more recent approaches to such matters have been in terms of inter and intra-personal kinds of thinking (Mayer 1989; Gardner 1993; Damasio 1994; LeDoux 1996; Damasio 1999).

The two terms of Problem solving and Critical thinking are commonly used as ways of differentiating kinds of thinking. Problem solving is the most generalised of notions. It usually seems to be more or less any kind of thinking, and there is not much reason for believing that because there is a problem to be solved there is a distinctive kind of thinking involved. Problem solving might be all kinds or any kind of thinking.

Critical thinking might be understood as a more or less distinctive kind of thinking in that it might be contrasted with formal and logical reasoning. Defined in this way, critical reasoning is the informal or plausible reasoning (distinguished from formal reasoning) envisaged in the DomProC model.

We could equate informal and plausible reasoning with social cognition, in contrast with formal and abstract reasoning. On the other hand we could place formal reasoning in the context of mathematics and science (as does Gardner), and in turn we could place informal reasoning and social cognition in the context of the social sciences and the humanities.

Kinds or Styles of Thinking

Ideally, we would want to distinguish different kinds or styles of thinking, but the psychometric tradition has not offered much help in making such distinctions. It tends to contrast abstract and formal thinking with thinking about different kinds of things like social cognition. As a result, formal reasoning becomes ***g***.

With these considerations in mind, the options shown in table 44 seem to offer themselves. We can try to distinguish kinds of reasoning. In this case, the distinction between formal and logical thinking and informal or critical reasoning offers itself for consideration. Or we could try to differentiate thinking in terms of content and topic, so that social cognition (thinking about the socio-cultural world) might be distinguished from thinking about the natural, material or mechanical world. This distinction might be most readily understood as the difference between thinking in mathematics and sciences and thinking in the social sciences and humanities: these context, or broad domain, distinctions might be underpinned by different kinds or styles of thinking.

These considerations suggest a model of differential abilities based on kinds of thinking or on kinds of broad domain.

Abilities as Kinds of Thinking and Broad Domains

The differentiation in terms of broad domains can and would benefit from incorporating different kinds of thinking, but differentiation in terms of kinds of reasoning does not readily distinguish between analytical/logical thinking and critical thinking. Hence is seems that the option 2 in Table 44 is the most comprehensive and the most likely to be meaningfully discriminant.

These considerations bring us back to the DomPro model developed at the end of Chapter 7. Like the review of the development of theories of abilities in Chapter 6, these considerations

suggest the strength of a model of performance and a definition of abilities that synthesises domains of knowledge and cognitive processes at a very high level of generality.

Table 44 Conventionally Testable Skills: Literacy and Kinds of Thinking about Broad Domains

Option 1 **Differential abilities as kinds of thinking**	Option 2 **Differential abilities as broad domains**
Literacy	Literacy
Analytical and logical thinking	Reasoning about scientific and technological issues (Scientific and technological understandings)
Critical thinking	Reasoning about cultural issues (Cultural understandings)
Social and interpersonal thinking	Reasoning about personal and interpersonal issues (Interpersonal understandings)

Some Overall Conclusions

This thesis has reviewed the weaknesses of psychometric approaches to ability, but it also shows strengths of psychometric methods and results. The psychometric approach can be narrow and uncritically certain, but it is also a powerfully focussed and disciplined approach that can clarify issues and dissolve some confusions. Significant limitations of the notion of abilities developed through the psychometric approach have been glanced at in different parts of this discussion, but, nonetheless, psychometric research offers some significant ideas about cognitive abilities and it should be seriously considered in developing notions of work-related abilities.

The strength of cognitive psychology and psychometrics is that they are concerned with thinking processes and kinds of thinking and are not limited to understanding abilities as kinds of tasks. The views of ability developed within cognitive psychology and psychometrics are generic because they are not particular tasks or knowledge and skills. These notions of generic abilities can usefully contribute to education and training, and they should inform conceptions of work-related skills.

Most of the work-related skills projects that have evolved in the last 15 to 20 years lack conceptual rigor because they were commonly pragmatic, consultative procedures that had little theoretical clarity or coherence. The pragmatism of these projects has meant they have not adequately addressed methodological matters or sought empirical grounding for the constructs they posit. On the other hand, DeSeCo set out to address basic matters of theory and procedure. It did not do so satisfactorily however because it was a reaction against the theories of ability developed within psychometrics and cognitive psychology, though it did not adequately deal with those theories.

In developing indicators of educational performance, the OECD is right to be concerned with wider issues than subject-based performances. However, in effectively ignoring much of psychometrics and cognitive psychology, DeSeCo has not built on, nor effectively eliminated from contention, ability theories based on cognitive psychology and psychometrics. It seems fair to conclude that DeSeCo actually lacks theoretical rigor, and as a result it is the least meaningful and practical of the projects on generic skills.

Furthermore, there is cause for concern about the influence of DeSeCo in Australia and about notions of employability skills that are being canvassed as superseding the Key Competencies. The proposals for employability skills being canvassed in Australia are not grounded in a coherent analysis or theory. Under the influence of DeSeCo, these proposals

take up the assessment of values and attitudes that the Mayer Committee treated very cautiously. These proposals lack a clear analysis of the implications of this move, and they do not present a coherent analysis or framework for dealing with such conceptions.

This thesis has offered an analysis of cognitive theory and research, an analysis of work-related and other abilities, and as a result formulated the DomPro model. The DomProAt model aims to be an overarching framework in which cognitive skills, personal attributes and attitudes and values can be comprehensively and coherently conceptualised.

The work reported here has also addressed the practical difficulties and challenges of assessing generic skills in secondary schools and shown that these difficulties are not insurmountable. The only realistic method of assessing generic skills in the school sector was sketched in the Mayer Committee assessment proposals. The whole-school assessment methodology conceptualised and implemented in the SBKCLAP trial showed that secondary teachers could in most cases make global, impression judgements of Key Competencies performances with no more than three minutes formal reflection per student. The judgements made by different teachers were quite consistent with each other, and as a result they can be validly and reliably used to develop an overall report for a student. The teachers participating in the trial judged the assessment procedures to be efficient and cost-effective.

Analysis of the assessments of teachers participating in the trial showed that pairs of teachers (from any combination of subject areas) typically produce an acceptable level of agreement in about 90% of cases. On three separate occasions further work on the whole-school assessment of Key Competences has demonstrated that the processes developed for SBKCLAP are practical and useful.

Whole-school assessment offers significant opportunities for secondary schools to assess and monitor kinds of performance that cannot be dealt with in subject-based assessments and to focus on the personal skills and personal growth of students. Teachers, students, parents and others value cross-curricular and generic skills (particularly the soft or personal skills like teamwork and planning and organising) and these can be formally assessed in whole school assessment. Cross-curricular and whole-school assessment can deal with kinds of soft or personal skills (the so called 'social outcomes' of schooling) that cannot be readily or reliably assessed by individual teachers.

The whole-school assessment process conceptualised in the SBKCLAP (and the WSA software that supports it) can help secondary schools answer the following questions.

- Which are the strongest and weakest competencies in our school?
- Which students are the highest and lowest on a particular competency?
- What are the comparative levels of performance of female and male students, or other sub-groups?
- How do individual teachers and the groups of teachers judge the performances of students on different competencies?
- To what extent do teachers agree with each other about different competencies?
- Do teachers have a consistent view of a competency?
- Have the views of teachers of the student competencies changed over time?
- Can we develop an assessable construct of our own?

The whole-school assessment process is a teacher assessment controlled by schools. It can only be a matter of choice whether schools adopt and implement it, even if it is encouraged and supported by central authorities. Schools may see whole-school assessment as a unique opportunity to strengthen the personal references that are currently produced by many secondary schools for exiting students.

An Epilogue

The Commonwealth Government invested 20 million dollars in the Key Competencies Pilot program between 1994 and 1996. The report on the outcomes of this investment made positive claims about the success of the program and the significance of the Key Competencies (MCEETYA 1996). Since 1996 the Key Competencies have been recognised and given favourable mention in various Australian curriculum documents. It is a requirement for accreditation that VET units and modules 'incorporate the Key Competencies'. The Key Competencies have had a significant influence on the national goals of schooling, but the integration of Key Competencies into the curriculum of general education and of VET training has seemed to add up to very little. No organization or system has proposed to assess the Key Competencies, and it seems that very little of the Mayer vision has been enacted or realised.

The report of the Key Competencies pilot phase called for further research, but the Commonwealth government funded hardly any Key Competencies projects after 1996, and the SBKCLAP was one of very few. The Commonwealth also funded a project managed by the National Industry Education Forum on portfolios. As part of that project, between 1998 and 2000, I further developed the software for whole-school assessment and reporting of Key Competencies conceptualised in the SBKCLAP .

With a change of Commonwealth government in 1996, the Key Competencies came to seem the initiative of a preceding government, and attention was given by the new government to notions of 'enterprise education' rather than Key Competencies. Between 1997 and 2000 it seemed that the Key Competencies (which it had been thought would take off like a wild fire) had quickly burnt out with a change of wind. Business and industry continued to support giving greater emphasis to generic, work-related abilities in education, and it might be the case that whatever direction business and industry faces the wind will eventually blow.

The National Centre for Vocational Education and Research has focused a significant part of its research program on generic skills. This program gave rise to the Australia National Training Authority funded paper on generic skills referred, to earlier in this chapter (Kearns 2001). The review by Curtis (Curtis 2001) also reported to the Business Council of Australia and the Australian Chamber of Commerce and Industry in 2001. Both of these reports made extensive reference to DeSeCo and other on-going international work on key competencies and work-related generic skills.

The review I undertook for the Victorian Board of Studies in 1996 (funded under the Commonwealth Key Competencies program) seemed to have no impact whatever when it was presented to the Board. At about that time, the Board was moving towards a substantial overhaul of the controversial Victorian Certificate of Education (VBOS 1999). As part of a review of the VCE the following proposal was developed:

> **Recommendation 19**
>
> That the Board of Studies develops appropriate means for reporting on key competencies and enterprise skills, informed by continuing Victorian, national and international work in this area.

The review noted that Victoria had been involved in assessing generic attributes through the Student Profile and it proposed to revive and enhance those procedures for the assessment of Key Competencies.

Current VCE	Enhanced VCE
KEY COMPETENCIES The Board of Studies provides schools with a template for a profile of student achievement on six generic attributes. The Board initially provided advice and guidelines on procedures for using the profiles, but these have not been reviewed or revised since 1993.	The Board of Studies will develop appropriate means for assessing and reporting on key competencies and enterprise skills, informed by continuing national and international work in this area. There will be a stronger emphasis on identifying key competencies within studies as part of the accreditation process.

In 2000 VBOS sought a consultant to review the relationship of the revised Curriculum and Standards Framework, VCE outcomes and VCE/VET units to the Key Competencies. In undertaking that work for the Board, I recommended the development of a method of whole-school assessment of the Key Competencies based on the SBKCLAP procedures described in Chapter 9. With the encouragement of the Board, in December 2000 I organised a trial of the software for whole-school assessment and reporting in ten Victorian schools. In June 2001 the VCAA (the new incarnation of the VBOS) authorised a field trial of my approach to whole-school assessment. The field trial was based on the whole-school assessment and reporting process and the software originally conceptualised in the SBKCLAP and refined in the NIEF Portfolio Project. Because the field trial of 2002 was judged to be an overall success, the VCAA has chosen to further develop the whole-school assessment and reporting process and software as a recommended method of generic skills and Key Competencies assessment for Victorian schools. There was a further trial of my method of whole-school assessment in 40 schools in 2003, and the VCAA has given further support to extend this work with some 60 schools in 2004.

In 2003 and 2004 there is a major review of the curriculum in Victoria that seems to move away from the Curriculum and Standards framework developed in the early 1990s. The review aims to identify Essential Learnings through Key Discipline Concepts and Skills that would circumscribe and concentrate the curriculum expectations placed on schools. While reducing the discipline-based aspect of the curriculum, the reform proposes that Essential Learnings include the following Generic Skills, Values and Attributes (VCAA 2004):

- communication;
- cognitive and metacognitive skills;
- social and Cultural skills;
- organizational and employability skills; and
- information and communication technology skills.

The story continues.

REFERENCES

ACACA (1995). ACACA Key Competencies Project Report. Canberra, ACACA.

ACCI and BCA (2002). Employability Skills for the Future. Canberra, Australian Chamber of Commerce and Industry, the Business Council of Australia and the Commonwealth of Australia.

ACTU (1993). Young People and Tomorrow's Workplace Strategy. 1993, http://www.actu.asn.au/research/policy/93youngp.htm.

Anastasi, A. (1988). Psychological Testing. London, Macmillan.

Anderson, M. (1992). Intelligence and Development: A Cognitive Theory. Oxford, Blackwell.

Barnett, R. (1994). The Limits of Competence. Knowledge, Higher Education and Society. Bristol, Open University Press.

Binkley, M., R. J. Sternberg, et al. (2000). An Overarching Framework for Understanding and Assessing Life Skills, Adult Literacy and Lifeskills Project http://www.ets.org/all.

Bloom, B. J., J. T. Hastings, et al. (1971). Handbook fo Formative and Summative Evaluation of Student Learning. New York, McGraw-Hill.

Bloom, B. S., D. R. Krathwohl, et al. (1956). Taxonomy of Educational Objectives. London, Longman, Green and Co.

Borthwick, A. (1992). "A Short Tour of the Findings of the Mayer Committee." Curriculum Perspectives (November 1992).

Borthwick, A. (1993). Key Competencies: Uncovering the Bridge between General and Vocational. Competencies: The Competencies Debate in Australian Education. C. Collins. Canberra, Australian College of Education.

Bowden, J. and G. N. Masters (1993). Implications for Higher Education of a Competency Based Approach to Education and Training. Canberra, AGPS.

Bryce, J. and A. Harvey Beavis (1997). Arts Education and the World of Work. Melbourne, ACER.

Bullock, A., O. Stallybrass, et al. (1988). The Fontana Dictionary of Modern Thought. London, Fontana Press.

Burt, C. (1940). The Factors of the Mind: An Introduction to Factor Analysis in Psychology. London, London University Press.

Calvin, W. H. (2001). How Brains Think: The Evolution of Intelligence. London, Weidenfeld and Nicolson.

Carmichael, L. (1992). Australian Vocational Certificate Training System. Canberra, National Board of Education and Training.

Carmichael, L. (1993). Workplace Imperatives for Education and Competence. Competencies. The Competencies Debates in Australian Education and Training. C. Collins. ACT, The Australian College of Education.

Carroll, J. B. (1982). The Measurement of Intelligence. Handbook of Human Intelligence. R. J. Sternberg. New York:, Cambridge University Press.: (pp. 29-120).

Carroll, J. B. (1993). Human Cognitive Abilities: A Survey of Factor-analytic Studies. New York, Cambridge University Press.

Carroll, J. B. (1995). "Reflections on Stephen Jay Gould's The Mismeasure of Man (1981): A retrospective review." Intelligence **21**: 121-134.

Carson, A. D. (2001). Research on Multiple Intelligences, http://www.geocities.com/Athens/Column/7568/research.html.

Cattell, R. B. (1987). Intelligence: Its Structure, Growth, and Action. New York, North-Holland.

CBC (1992). Employability Skills Profile: What are Employers Looking For? Ottawa, Conference Board of Canada.

CBC (2000). Employability Skills 2000+. Ottawa, Conference Board of Canada.

CBI (1989). Towards a Skill Revolution. London, Confederation of British Industry.

Ceci, S. J. (1990). On Intelligence...More or Less. A Bio-ecological Treatise on Intellectual Development. Englewood Cliffs, N.J., Prentice Hall.

Ceci, S. J., S. L. Ramey, et al. (1990). "Framing Intellectual Assessment in Terms of a Person-process-context Model." Educational Psychologist **25**: 269-291.

Chomsky, N. (1980). Rules and Representations. New York, Colombia University Press.

Chomsky, N. (1986). Knowledge of Language: Its Nature, Origin, and Use. New York, Praeger.

Collier, G. (1994). Social Origins of Mental Ability. New York, John Wiley and Sons.

Collins, C. (1993). Competencies: The Competencies Debate in Australian Education and Training. Canberra, Australian College of Education.

Collins, C. (1996). "What Teachers Need to Know? The Competencies Debate." South Australian Educational Leader 7(Nos 3).

Craigie, D. (1995). How Could the 'Generic Skillers' and the 'Competencies Buffs' be Both Right and Wrong? A Defence of Organised Knowledge. Restructuring the Curriculum: Choosing the Future, Belconnen ACT, Australian Curriculum Studies Association.

Curtis, P. M. D. (2001). Employability skills for Australian industry: literature review and framework development. Melbourne, ACER.

Damasio, A. (1994). Descartes' Error. New York, Grosset/Putnam.

Damasio, A. (1999). The Feeling of What Happens. Body and Emotion in the Making of Consciousness. London, William Heinemann.

Das, J. P., J. Kirby, et al. (1975). "Simultaneous and Successive Synthesis: An Alternate Model for Cognitive Abilities." Psychological Bulletin **82**.

DEETYA (1994). Employers Groups and Unions Support the Key Competencies. Canberra, DEETYA.

DEETYA (1994). Key Competencies' Endorsements. Canberra, DEETYA.

Ducker, C. (1993). Training Reform and Key Competencies. Competencies: The Competencies Debate in Australian Education and Training. C. Collins. Canberra, Australian College of Education.

Edwards, B. and J. Smith (1996). Issues and Options in the Assessment and Reporting of Key Competencies in the School Sector. Canberra, The Australian Curriculum, Assessment and Certification Authorities (ACACA).

Finn, B. (1992). Young People's Participation in Post-Compulsory Education and Training. Canberra, Australian Education Council Review Committee.

Flew, A. (1983). A Dictionary of Philosophy. London, Macmillan Press.

Fodor, J. A. (1983). The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, Massachusetts, MIT/Bradford Press.

G.Kearsley (1994). Howard Gardner. The Theory in Practice Database, http://tip.psychology.org/index.html.

Gagne, E. D., C. W. Yekovich, et al. (2002). The Cognitive Psychology of School Learning. New York, Longman.

Gardner, H. (1973). The Arts and Human Development. New York, Basic Books.

Gardner, H. (1974). The Quest for Mind. Piaget, Levi Strauss and the Structuralist Movement. New York, Vintage Books.

Gardner, H. (1974). The Shattered Mind. The Person After Brain Damage. New York, Vintage Books.

Gardner, H. (1985). Frames of Mind: The Theory of Multiple Intelligences. New York, Basic Books.

Gardner, H. (1987). The Mind's New Science: A History of the Cognitive Revolution. New York., Basic Books.

Gardner, H. (1991). The Unschooled Mind: How Children Think and How Schools Should Teach. New York, Basic Books.

Gardner, H. (1992). Assessment in Context: The Alternative to Standardized Testing.

Gardner, H. (1993). Frames of Mind: The Theory of Multiple Intelligences. London, Fontana Press.

Gardner, H. (1993). Reflections on Multiple Intelligences: Myths and Messages, http://www.byu.edu/pe/blakemore/reflections.html.

Gardner, H. (1999). The Disciplined Mind: What All Students Should Understand. New York, Simon and Schuster.

Gardner, H., M. L. Kornharber, et al. (1995). Intelligence: Multiple Prespectives. Harvard, Harvard University Press.

Goleman, D. (1995). Emotional Intelligence: Why It Can Matter More Than IQ. New York, Bantam Books.

Gonczi, A., P. Hager, et al. (1993). The Development of Competency-based Assessment Strategies for the Professions. Canberra, National Office of Overseas Skill Recognition.

Gonczi, A., P. Hager, et al. (1992). A Guide for the Development of Competency Standards of the Professions. Research paper 7. ACT, Australia. Dept. of Employment Education and Training, Australia. National Office of Overseas Skills Recognition.

Gonczi, A., P. Hager, et al. (1990). Establishing Competency-Based Standards in the Professions. Canberra, National Office of Overseas Skill Recognition.

Gottfredson, L. (1997). "Why *g* Matters: The Complexity of Everyday Life." Intelligence(24).

Gottfredson, L. S. (1998). The General Intelligence Factor. Scientific American. **9**.

Gould, S. J. (1981). The Mismeasure of Man. New York, Penguin Books.

Gould, S. J. (1996). The Mismeasure of Man (Revised). London, Penguin Books.

Greenfield, S. (1997). The Human Brain: A Guided Tour. London, Weidenfeld and Nicolson.

Guilford, J. P. and R. Hoepfner (1971). The Analysis of Intelligence. New York,, McGraw-Hill.

Hearnshaw, L. S. (1979). Cyril Burt Psychologist. London, Hodder and Stoughton.

Herrnstein, R. J. and C. A. Murray (1994). The Bell Curve: Intelligence and Class Structure in American life. New York, Free Press.

Hill, P. (1995). Student Assessment and the Key Competencies, Centre for Applied Educational Research, Faculty of Education, The University of Melbourne.

Horn, J. L. (1994). Fluid and Crystallized Intelligence. The Encyclopaedia of Human Intelligence. R. J. Sternberg. New York, Macmillan. **Vol. 1:** 443-451.

Hunt, E. B. (1978). "Mechanics of Verbal Ability." Psychology Review **85**: 109-130.

Hunt, E. B., C. Lunneborg, et al. (1975). "What Does It Mean to Be High Verbal?" Cognitive Psychology **7**: 194-227.

Hunt, M. (1993). The Story of Psychology. New York, Doubleday.

Jensen, A. R. (1980). Bias in Mental Testing. New York, Free Press.

Jensen, A. R. (1998). The *g* Factor. Westport, Praeger.

Kearns, P. (2001). Generic Skills for the New Economy: A Review of Research. Adelaide, NCVER.

Kearsley, G. (1994). Howard Gardner. The Theory in Practice Database, http://tip.psychology.org/gardner.html.

Kline, P. (1991). Intelligence: The Psychometric View. New York:, Routledge.

Kline, P. (1998). The New Psychometrics. Science, Psychology and Measurement. London, Routledge.

Lave, J. (1988). Cognition in Practice: Mind, Mathematics and Culture in Everyday Life. New York, Cambridge University Press.

LeDoux, J. (1996). The Emotional Brain. The Mysterious Underpinnings of Emotional Life. London, Pheonix.

Lilly, M. (1996). Integrating Key Competencies. Melbourne, Australian Competency Research Centre.

Lilly, M. (1997). Did Mayer Get It Wrong?, Australian Training Review.

Lohrey, A. (1995). A Report on Transferability in Relation to the Key Competencies. Sydney, Centre for Workplace Communication and Culture UTS.

Marginson, S. (1992). "Education After Finn. Shifting the Emphasis." Independent Education **22**.

Masters, G. N. and D. McCurry (1990). Competency-based Assessment in the Professions. National Office of Overseas Skills Recognition. Canberra, AGPS.

Mayer, E. (1992). The Key Competencies. Melbourne, Australian Education Council and Ministers of Vocational Educational, Employment and Training.

Mayer, J. D., Caruso, D., Zigler, E., Dreyden, J. (1989). "Intelligence and Intelligence-related Personality Traits." **Intelligence, 13.**

Mayer, J. D. and P. Salovey (1997). What is Emotional Intelligence? Emotional Development and Emotional Intelligence. P. Salovey and D. J. Sluyter. New York, Basic Books.

McCurry, D. (1996a). Approaches to Assessing the Key Competencies. Melbourne, ACER.

McCurry, D. (1996b). Approaches to Assessing the Key Competencies in Victoria. Melbourne, ACER.

McCurry, D. (2001). An Audit of the Key Competencies in the CSF and the VCE Outcomes. Melbourne, ACER.

McCurry, D. and J. Bryce (1997). The School-based Key Competencies Levels Assessment Project: Final Report. Melbourne, ACER.

McCurry, D. and J. Bryce (2002). Report of the VCAA Whole School Assessment Trial of 2002. Melbourne, ACER.

MCEETYA (1996). Report on the Outcomes of the Key Competencies Pilot Phase. Canberra, Ministers for Education, Employment, Training and Youth Affairs.

MCEETYA (1999). Adelaide Declaration, National Goals of Schooling in the Twenty First Century. Adelaide, http://www.curriculum.edu.au/mceetya/nationalgoals/natgoals.htm.

McGaw, B. (1992). Competence or Expertise? The ACER Annual Report of 1992, ACER.

McGaw, B. (1996). Their Future: Options for Reform of the Higher School Certificate. Sydney, NSW DTEC.

Mellor, S., G. Withers, et al. (1995). Literacy and the Competencies: Teachers' Perspectives. Melbourne, ACER.

Messick, S. (1992). Multiple intelligences or Multilevel Intelligence?: Selective Emphasis on Distinctive Properties of Hierarchy : an Essay Review of Gardner's Frames of Mind. Research report ; RR-92-29. Princeton, N.J., Educational Testing Service.

Neisser, U., G. Boodoo, et al. (1996). "Intelligence: Knowns and Unknowns." American Psychologist, 51(2),

NIEF (1995). Business, Industry and Key Competencies. Melbourne, National Industry Education Forum.

NIEF (2000). NIEF Key Competencies Portfolio Project Manual. Melbourne, National Industry and Education Forum.

NLLIA (1994). Cultural Understandings as the Eighth Key Competency. Final Report, James Cook Univ, Townsville, Technology Univ , Sydney.

NZQA (1993). Essential Skills and Generic Skills in the National Qualifications Framework. Wellington, New Zealand Qualifications Authority.

OECD (1997). Prepared for Life? How to Measure Cross-curricular Competencies. Paris, OECD.

OECD (1999). Definition and Selection of Key Competencies. Paris, Organisation for Economic Cooperation and Development.

OECD (2001). Introduction to DeSeCo. Paris, Organisation for Economic Cooperation and Development http://www.statistik.admin.ch/stat_ch/ber15/deseco/deseco_backgrpaper_dec01.pdf.

OECD (2002). Understanding the Brain: Towards a New Learning Science. Paris, Organisation for Economic Cooperation and Development http://www1.oecd.org/publications/e-book/9102021e.pdf.

Phenix, P. H. (1964). Realms of meaning. New York, McGraw-Hill.

Piaget, J. (1971). The Psychology of Intelligence. London, Routledge & Kegan.

Piper, K. (1997). Riders in the Chariot. Curriculum Reform and the National Interest 1965-95. Melbourne, ACER.

QERC (1985). Quality of Education in Australia, Report of the Review Committee (Karmel Report). Canberra, AGPS.

Reber, A. S. (1985). The Penguin Dictionary of Psychology. London, Penguin Books.

Rowe, H. A. H. (1985). Problem Solving and Intelligence., Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumsey, D. and B. Hannan (1996). Discussion Paper on the Redrafting of the Proposed Eighth Key Competency. Melbourne, Curriculum Corporation, Standards and Curriculum Council.

Ryan, C. (1997). NSW Key Competencies Pilot Project Report. Sydney, NSW Department of Training and Education Co-ordination.

Rychen, D. S. and L. H. Salganik (2001). A Contribution to the OECD Program Definition and Selection of Competencies: Theoretical and Conceptual Foundations. Seattle, Hogrefe & Huber.

Rychen, D. S. and L. H. Salganik, Eds. (2003). Key Competencies for a Successful and Well-functioning Society. Gottingen, Hogrefe & Huber.

SCAA (1996). Review of Qualifications for 16-19 Year Olds. London, School Curriculum and Assessment Authority.

SCANS (1992). The Secretary's Commission on Achieving Necessary Skills (SCANS) Final Report. Washington DC, US Department of Labour.

Spearitt, D. (1996). "Carroll's Model of Cognitive Abilities: Educational Implications." International Journal of Educational Research **25**(2).

Spearman, C. (1927). The Abilities of Man: Their Nature and Measurement. London, Macmillan and co.

Spearman, C. (1946). "Theory of the general factor." British Journal of Psychology **36**:

Stanley, G. (1993). The Psychology of Competency-based Education. Competencies. The Competencies Debates in Australian Education and Training. C. Collins. ACT, The Australian College of Education.

Stanton, J. (1995). Business, Industry and Key Competencies. Melbourne, National Industry Education Forum.

Sternberg, R. J. (1977). Intelligence, Information Processing and Analogical Reasoning: the Componential Analysis of Human Abilities. Hilesdale, NJ, Erlbaum.

Sternberg, R. J. (1983). How Can We Teach Intelligence? Philadelphia, PA, Research for Better Schools.

Sternberg, R. J. (1985). Beyond IQ: A Triarchic Theory of Human Intelligence. New York, Cambridge University Press.

Sternberg, R. J. (1985). Human Abilities : an Information-processing Approach. New York, W.H. Freeman.

Sternberg, R. J. (1989). The Triarchic Mind: A New Theory of Human Intelligence. New York, Penguin Books.

Sternberg, R. J. (1990). Metaphors of Mind : Conceptions of the Nature of Intelligence. Cambridge ; New York, Cambridge University Press.

Sternberg, R. J. (1991). Theory-based Testing of Intellectual Abilities. Intelligence: Reconceptualization and Measurement. H. Rowe.

Sternberg, R. J. (1994). Encyclopaedia of Human Intelligence. New York, Macmillan.

Sternberg, R. J. (1996). IQ Counts, but What Really Counts Is Successful Intelligence. National Association of Secondary School Principals Bulletin, National Association of Secondary School Principals.

Sternberg, R. J. (1996). Successful Intelligence : How Practical and Creative Intelligence Determine Success in Life. New York, Simon & Schuster.

Sternberg, R. J. (1997). Thinking Styles. Cambridge, Cambridge University Press.

Sternberg, R. J. (2000). The Handbook of Intelligence. Cambridge, Cambridge University Press.

Thurstone, L. L. (1938). Primary Mental Abilities. Chicago, University of Chicago Press.

VBOS (1999). Report of the Committee of Review on the Victorian Certificate of Education. Melbourne, Victorian Board of Studies.

VCAA (2004). Victorian Curriculum Reform 2004: A Framework for Essential Learnings. Melbourne, Victorian Curriculum and Assessment Authority.

VCAB (1990). VCE Assessment Policy. Melbourne, Victorian Curriculum and Assessment Board.

VCAB (1994). The VCE Student Profile. Melbourne, Victorian Curriculum and Assessment Board.

Vernon, P. E. (1973). The Structure of Human Abilities. London, Methuen.

Vogt, W. P. (1993). Dictionary of Statistics and Methodology. New York, Sage Publications.

Vygotsky, L. (1986). Thought and Language. Cambridge Massachusetts., The MIT Press.

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Cambridge M. A., Harvard University Press.

WCDE (1997). Outcomes Education in South Africa. Background information for Educators. Capetown, Western Cape Department of Education.

Weinert, F. E. (1999). Concepts of Competence: DeSeCo Expert Report. Neuchâtel, Max-Planck Institute for Psychological Research, Swiss Federal Statistical Office.

Weinert, F. E. (2001). Competencies: A Conceptual Clarification. Defining and Selecting Key Competencies. L. H. Salganik and D. S. Rychen. Seattle, Hogrefe & Huber.

Wertsch, J. V. (1985). Culture, Communication and Cognition: Vygotskian Perspectives. Cambridge Mass., Cambridge University Press.

Wolf, A. (1991). "Assessing Core Skills: Wisdom or Wild Goose Chase?" Cambridge Journal of Education **21**(2).

Wolf, A. (1995). Competence-based Assessment. Buckingham England, Open University Press.

Wolfe, P. (2001). Brain Matters: Translating Research into Classroom Practice. Alexandria, Association of Supervision and Curriculum Development.

Wozniak, R. H. (1999). Classics in Psychology, 1855–1914: Historical Essays. Chicago, Thoemmes Press.

Appendices

Appendix 1
The National Goals of Schooling 1999

The Adelaide Declaration on National Goals for Schooling in the Twenty-First Century, Ministerial Council on Education, Employment, Training and Youth Affairs, 1999.

Schooling should develop fully the talents and capacities of all students. In particular, when students leave schools they should

1 have the capacity for, and skills in, analysis and problem solving and the ability to communicate ideas and information, to plan and organise activities and to collaborate with others

2 have qualities of self-confidence, optimism, high self-esteem, and a commitment to personal excellence as a basis for their potential life roles as family, community and workforce members

3 have the capacity to exercise judgement and responsibility in matters of morality, ethics and social justice, and the capacity to make sense of their world, to think about how things got to be the way they are, to make rational and informed decisions about their own lives and to accept responsibility for their own actions

4 be active and informed citizens with an understanding and appreciation of Australia's system of government and civic life

5 have employment related skills and an understanding of the work environment, career options and pathways as a foundation for, and positive attitudes towards, vocational education and training, further education, employment and life-long learning

6 be confident, creative and productive users of new technologies, particularly information and communication technologies, and understand the impact of those technologies on society

7 have an understanding of, and concern for, stewardship of the natural environment, and the knowledge and skills to contribute to ecologically sustainable development

8 have the knowledge, skills and attitudes necessary to establish and maintain a healthy lifestyle, and for the creative and satisfying use of leisure time

Appendix 2
Principles for the Reporting on Individuals' Key Competency Achievements from the Mayer Report (p.xvii)

The Mayer Committee offered the following principles for assessing and reporting on individuals' performance.

The Committee recommends that:

The AEC and MOVEET agree in principle to the following set of principles for reporting on individual achievement of the Key Competencies:

- An individual record of performance showing achievement in the Key Competencies not previously reported should be made available to students and trainees at any point at which they exit from a school program at Years 11–12 level or from a recognised entry-level program of vocational education and training.
- Assessments reported on records of performance should be in accordance with nationally-agreed principles for assessment.
- Records of performance should be based on a common format.
- The issuing of records of performance should be subject to consistent provisions regarding confidentiality and the maintenance of databases of records.

The Committee also offered the following principles for a national reporting of Key Competencies performance.

Principles for the National Reporting of Key Competency Achievements from the Mayer Report (p.xvii)

The Committee recommends that:

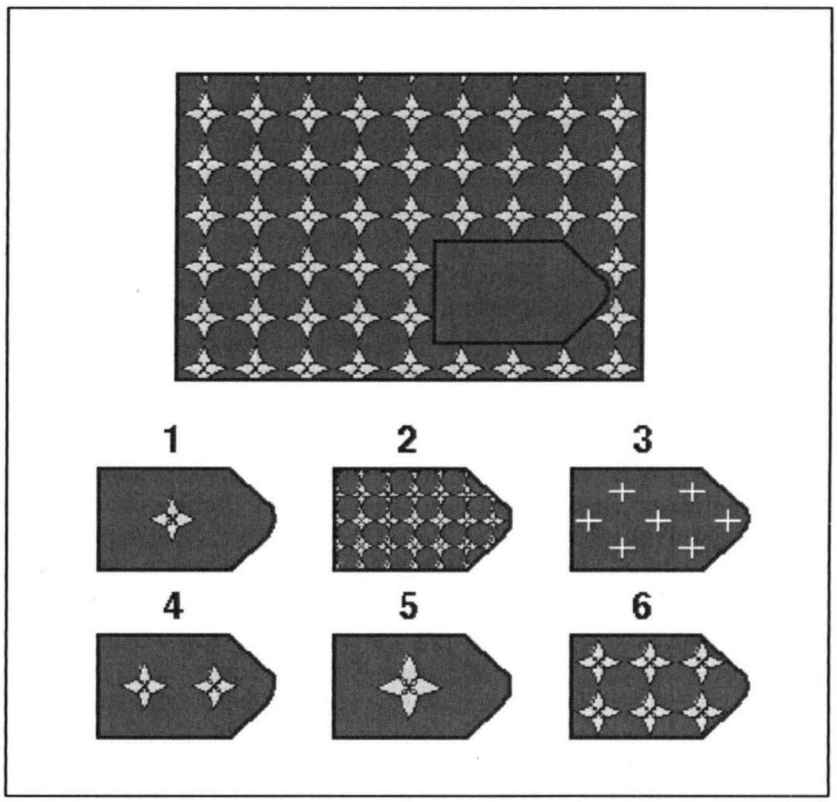
The AEC and MOVEET agree in principle to the following set of principles for national reporting on achievement of the Key Competencies:

- The collection of data for the purpose of national reporting should be designed to provide a publicly credible means of reporting on the extent to which the Key Competencies are being achieved, to improve the information base for evaluating the effectiveness of education and training programs, and contribute to meeting the needs of public accountability.
- The collection of data for the purpose of National reporting should be undertaken on the basis of sample statistics drawn from databases of individual records of performance.
- Statistical sampling techniques should provide for analyses of participation and achievement by access and equity target groups.
- The collection of data for the purpose of national reporting should incorporate procedures to protect the confidentiality of individuals and the identity of individual schools and training providers.

Appendix 3
Raven's Progressive Matrices

The items in Raven's Progressive Matrices are the archetypical fluid or abstract reasoning tasks.

The items involve abstract symbols and reasoning about visual patterns and relationships. It is argued by some that such test items are the best test of intelligence in that they are not curriculum based and they are non-verbal. Such abstract, fluid reasoning contrasts markedly with thinking in language and through language.



Appendix 4
An Overview of the Assessment and Reporting Components of Key Competency Projects in the Various States and Territories

This work on Key Competencies assessment and reporting undertaken in the pilot phase showed an awkward and inconsistent set of activities in the different states and systems.

The ACT Department of Education and Training considered ways of integrating Key Competency assessment and reporting into existing practices. This work explored the feasibility of including the assessment and reporting of Key Competencies as a sub-set of subject-based assessment and certification procedures so that attainment of certain levels of performance on the subject studies would be translated into a Key Competencies level of performance.

Similarly the Tasmanian Department of Education and the Arts explored the integration of Key Competencies assessment into the Tasmanian Certificate of Education. The Key Competencies were embedded in syllabuses, and hence this seemed to offer the opportunity for Key Competencies reporting through the subject based assessment and reporting. It was thought that this arrangement could mean that the Tasmanian Certificate of Education grade related reporting of attainment on criteria could also involve the assessment and reporting of Key Competencies.

The Board of Senior Secondary School Studies in Queensland explored the possibility of determining performance levels for at least 5 Key Competencies from the Queensland Core Skills Test and subject grades as the means of reporting Key Competencies levels on the Senior certificate. The Board also explored the possibility of school issued descriptive assessments of two Key Competencies that are outside the formal curriculum. The Queensland Department of Education focused on assessing and reporting Key Competency performance in the 'complementary curriculum' and work placements. This work examined the use of a school report or a Student Key Competencies Portfolio as a means of reporting Key Competency assessments. It was envisaged that the portfolio could include a summative assessment of Key Competency levels, verbal descriptors from a standardised and codeable bank, a log of Key Competency related work kept by the student, work samples and other documentation.

The Victorian Board of Studies considered two options for Key Competencies assessment and reporting. The first would be solely school-based, and the second would involve Board monitored and endorsed school assessments. The second option would involve Key Competencies assessment and reporting on the VCE certificate. The following possibilities were considered for such reporting:

- a single statement on the VCE certificate that the Key Competencies have been achieved;
- a statement reporting levels of achievement for each Key Competency;
- a range of pre-formulated descriptive comments; and
- teachers' open-ended comments.

In Western Australia the Secondary Education Authority explored the embedding of Key Competencies as an integral part of some courses, and the use of a common assessment framework approach as a basis for Key Competencies assessment. The Secondary Education Authority identified the following assessment alternatives:

- assessment and reporting by schools: no certification by the Secondary Education Authority;

- certification of one level of achievement of the Key Competencies on the Year 12 certificate;

- certification of the highest of three levels of achievement in each of the Key Competencies; and

- certification of three levels of achievement of the Key Competencies, based on an aggregated score.

The New South Wales Board of Studies explored a range of assessment models and appropriate reporting structures for Key Competencies, including teacher judgements and common assessment tasks.

The Northern Territory Board of Studies explored the use of descriptive profiles generated by teachers, supported by the development of school context pointers. The Board has also considered Key Competencies assessment through school curriculum vitae or personal record keeping as the basis for school-based exit references.

The Key Competencies project in South Australia of the Ministry and the Board of Studies did not accept the 3 levels of performance proposed by the Mayer Committee because it did not seem feasible or affordable to make valid, consistent and reliable judgements across more than one level for seven Key Competencies. The aim of the assessment and reporting work in South Australia was to set a single benchmark for each Key Competency as a post-compulsory standard and entry-level requirement. The intention was to produce a set of Key Competency descriptors as the basis for school derived grades which could be incorporated in the South Australian Certificate of Education.

A number of things are evident from this summary.

- Very little of the assessment work involved the kind of direct assessment of Key Competencies by teachers envisaged in the Mayer Report.

- Some systems envisaged integrating the Key Competencies assessment with subject assessment. ACT, Tasmania and Queensland envisaged inferring Key Competencies assessments from subject assessments or tests.

- There was some openly declared opposition to assessment of levels of performance on the Key Competencies.

- The involvement of the boards of studies in any system was not uniformly embraced.

- There was little coherence or consistency of approach in the different proposals.

National consistency of approach had certainly not resulted from the state based Key Competencies projects, and there was considerable doubt about the validity, reliability and feasibility of the kind of provider based assessment of Key Competencies envisaged by the Mayer Committee in the various boards of studies.

Appendix 5
The SBKCLAP Teacher Background Notes and Assessment Framework

PARTICIPANT BACKGROUND NOTES

THE KEY COMPETENCIES

LEVELS ASSESSMENT

TRIAL

A DEETYA/ACER PROJECT

EXPLORING THE SCHOOL-BASED ASSESSMENT

OF KEY COMPETENCIES LEVELS FOR GRADES 11 & 12

August 1996

THE KEY COMPETENCIES LEVELS ASSESSMENT TRIAL

These notes are an introduction for participating teachers to a project by a team of researchers from the Australian Council for Educational Research (ACER) funded by the Department of Employment, Education, Training and Youth Affairs (DEETYA) on the school-based assessment of the Key Competencies.

These notes:

- give some background about the development of the Mayer Key Competencies;
- outline an adaptation made of the Key Competencies for school-based assessment in this project; and
- describe the assessment procedure to be trialed in the project.

The ACER team and DEETYA are grateful to the schools and teachers who are participating in the trial, and every effort will be made to make efficient use of the time and intellectual energy of busy professionals. The ACER team and DEETYA believe this trial will produce valuable information that will make an important contribution to the further development of the Key Competencies.

1 The Development of the Key Competencies

The Key Competencies are the major development in the assessment of school and vocational education of the 1990s. They were developed as part of a concern to bring general and vocational education closer together, and to better prepare young people for the world of work.

In 1990 the Australian Education Council commissioned a review of post-compulsory education and training. The Finn Committee, as it was known, produced a Report entitled *Young People's Participation in Post-Compulsory Education and Training*. Among other things, this report concluded that there are certain things or 'competencies' in a number of areas that all young people need to learn in their preparation for employment.

THE FINN KEY AREAS

Language and communication

Using mathematics

Scientific and technological understanding

Cultural understanding

Problem solving

Personal and interpersonal

In 1991 a Committee chaired by Eric Mayer was charged by the Australian Education Council with identifying the key competencies that underpin success in employment. The purpose of these competencies would be to provide a common reference point for curriculum and teaching in both school and training sectors, and to provide the basis for a nationally consistent approach to assessing and reporting achievement.

MAYER KEY COMPETENCY STRANDS
Collecting, analysing and organising ideas and information
Communicating ideas and information
Planning and organising activities
Working with others and in teams
Using mathematical ideas and techniques
Solving problems
Using technology
Cultural understanding

The Report of the Mayer committee, entitled *The Key Competencies: Putting General Education of Work*, stated that 'there are certain essential things that all young people need to learn in their preparation for employment'. It was also stated that students should be able to develop these employment-related Key Competencies regardless of the pathway they follow through education and training in the post-compulsory years. The Report envisaged Records of Performance on Key Competencies that would be issued to individuals as information for potential employers and for the purpose of credit transfer between different educational institutions.

Between 1994 and 1996 DEETYA mounted a program of trialing the Key Competencies in schools and vocational education. This trialing has shown that the Key Competencies are commonly implicit and sometimes explicit in the curriculum of general education. This trialing has also shown that the Key Competencies have been a valuable means of developing teaching and learning. Comparatively little thought has been given, on the other hand, to the assessment of Key Competencies in general education, and this project is an attempt to address this issue.

2 An ACER/DEETYA trial of a method for the school-based assessment and reporting of Key Competencies levels

Since the publication of the Key Competencies a number of concerns and reservations have been expressed about the Key Competencies and the possibilities for assessing them in the school system. These concerns will not be outlined here, but this project is a response to some to them, and an attempt to develop a valid, reliable and cost effective assessment of Key Competencies levels in general education programs.

This trial is based on the views that: (1) the Key Competencies can be meaningfully and realistically assessed in general education programs; and (2) that they can be assessed, in a minimal but useful fashion, without changes in current school programs, and without placing substantial new burdens on teachers and students.

This project deals only with the assessment of Key competencies, and, in assessment, it only deals with the reporting of levels of performance. It does not address the issues of teaching and learning which are commonly viewed as the most valuable stimulus the Key Competencies have to offer education systems. The scheme trialed in this project is minimal, and it does not claim to do justice to the opportunities offered by the Key Competencies for progress in general education.

This trial is designed to test the following propositions:

1 Although in some respects the Key Competencies are unfamiliar to teachers, they can make sense to teachers, and they can be assessed globally by teachers.

2 The Key Competencies are part of general education; teachers know things about the Key Competencies performance of their students; and the Key Competencies can be assessed as part of current school programs.

3 Key Competencies assessments are not in competition with or a replacement for subject-based assessments. They offer information that is in addition to that offered by subject-based assessments.

4 Key Competencies assessments are aimed particularly at offering information about students who are in transition to work, and they touch on skills and attitudes not always or readily included in subject assessments.

5 The most realistic way of reporting summary judgements of Key Competencies performance in general education is on the basis of global judgements made by individual teachers about the general skills and attitudes of the students they teach, with these individual judgements being subsequently synthesised into and reported as overall school levels.

6 Judgements about Key Competencies performance should be broad and general inferences about the way students are likely to perform in the future rather than judgements about students' performances in specific courses. Because they are general or generic, Key Competencies judgements can be synthesised into collective school judgements.

The challenge in developing the kind of school-based assessment of Key Competencies envisaged by this project is to design a scheme that can be practically and efficiently implemented by teachers, and that can offer valid and reliable information about students' performance.

This trial is an attempt to develop and test a practical, reliable and valid scheme of school-based assessment of the Key Competencies.

The following page presents an outline of the Key Competencies as adapted for this project.

The Key Competencies for Assessment of Grades 11 and 12 in General Education

(adapted from the Mayer Key Competencies changes or additions are *italicised*)

The Key Competencies, as understood in this project, are a mixture of knowledge, skills *and attitudes* that all young people need to enable them to participate effectively in the emerging forms of work and work organisation.

1 Collecting, Analysing and Organising Information

The capacity to locate information, sift and sort information in order to select what is required and present it in a useful way, and evaluate both the information itself and the sources and methods used to obtain it.

2 Communicating Ideas and Information in *Speech and Writing*

The capacity to communicate effectively with others *using speech and writing.*

3 Using Mathematical Ideas and Techniques

The capacity to use mathematical ideas, such as number and space, and techniques, such as estimation and approximation, for practical purposes.

4 Using Cultural Understandings

The capacity to understand and interpret socio-cultural ideas, and to reason plausibly about personal and inter-personal issues.

5 Solving Problems

The capacity *and the attitude* needed to apply problem-solving strategies in purposeful ways, both in situations where the problem and the desired solution are clearly evident and in situations requiring critical thinking and a creative approach to achieve an outcome.

6 Using Technology

The capacity *and the attitude* needed to apply technology, combining the physical and sensory skills needed to operate equipment with the understanding of scientific and technological principles needed to explore and adapt systems.

7 Planning and Organising Activities

The capacity *and the attitude* needed to plan and organise one's own work activities, including making good use of time and resources, sorting out priorities and monitoring one's own performance.

8 Working with Others and in Teams

The capacity *and the attitude* needed to interact effectively with other people both on a one-to-one basis and in groups, including understanding and responding to the needs of a client and working effectively as a member of a team to achieve a shared goal.

3 The trial procedure

In this project a number of schools are being asked to be part of a 'cold start and minimal trial' of a procedure for assessing Key Competencies in general education programs. It is a cold start and minimal trial because participating schools know comparatively little about the Key Competencies, and comparatively little support will be offered to the schools to make the assessment. The belief to be tested in the project is that the teachers in the trial will be able to make sense of the Key Competencies, and they can use them to report meaningful information about their students.

In this project groups of teachers who teach a common group of Year 11 students in a number of trial schools will be given a brief introduction to the Key Competencies and the trial assessment procedures by members of the project team in August 1996. The teacher participants will also be given a **KC Facets and Standards Schema**, and the **KC Facet Descriptors Overview** to use in making the assessments.

The teachers in the trial will be asked to determine Key Competencies levels for a common group of students, and these assessments will be synthesised into an overall school judgement using specified procedures.

The following discussion outlines the interpretation of the Key Competencies made for this trial, and the procedures to be used in making the individual teacher judgements and the overall school assessment during the trial.

4 The KC adaptation for this trial

The KC outline above is the Mayer conception of the Key Competencies, except that they have been re-ordered and have been adapted to make explicit an emphasis on attitudes as a clear component of some of the Key Competencies. Saying that a competency involves capacities and attitudes is to see it as involving what is usually thought of as knowledge and skills, and also more personal attributes or characteristics.

The Mayer Committee rejected the expectation that the Key Competencies involved the assessment of 'attitudes' (what they meant by the term is unclear). But there is reason for thinking that a number of the Key Competencies place an emphasis on attitudes that is not found in most formal educational assessments which emphasise knowledge and skills but do not assess more personal attributes. It has been decided to explicitly recognise the place of attitudes in the second four Key Competencies of this trial because it is believed it will increase the validity and the distinctiveness of the Key Competencies assessment, and it is believed it will improve the reliability of the assessment in general education programs.

There is a principled as well as a practical reason for moving this implementation of Key Competencies towards the assessment of attitudes as well as the assessment of knowledge and thinking skills.

One of the major difficulties in designing a practical Key Competencies assessment in schools has been how to envisage the assessment of Planning and organising activities, Working with others and in teams and Using technology when the more specific activities involved in these Key Competencies are not necessarily required, or, in a strict sense, are not assessed in current programs. But this problem can be side stepped and the essential features of the Key Competencies emphasised if the assessment is envisaged as general kinds of approaches or attitudes rather than specific performances.

On the other hand, Solving problems has been criticised as so general as to be meaningless, but this notion can be given meaning (and meaning that respects the spirit that nominated it

as important in the first place) if it is seen as involving attitudes as well as knowledge and skills. Explicitly recognising the element of attitudes in the more problematic Key Competencies makes the focus and distinctiveness of these Key Competencies clearer, and they should then be easier to operate as an assessment.

Using Cultural Understandings is the most controversial of the Key Competencies. In some respects it has been controversial because Using Cultural Understandings seems vague to some critics, and it has seemed a matter of socio-political value judgements to others. Using Cultural Understandings is included here because it is thought to cover an important and fairly specific kind of thinking (thinking about inter- and intra-personal and socio-cultural issues), and it is placed among the first four Key Competencies that emphasise thinking skills rather than those that include attitudes to avoid the danger of drifting into assessments of values or ideology.

The principled reasoning for explicitly introducing attitudes into the Key Competencies assessment is to recognise that competence in employment is usually defined as involving knowledge, skills and attitudes.

5 The assessment of general skills and attitudes

As is the case with most conceptions of the goals of general education, the Key Competencies, as adapted for this trial, involve varying mixtures of knowledge, skills and attitudes. But the Key Competencies are concerned with that which is general to education (and might be called cross-curricular competencies) rather than the knowledge and skills that are more or less specific to individual subjects and subject-based assessments.

The Key Competencies are somewhat different in the things each emphasises. The first four Key Competencies emphasise kinds of thinking skills — the same sorts of thinking skills that are the basis of the school curriculum, or, more precisely, the skills assessed in school grades. The second four can also be thought of as thinking skills, but the emphasis in the descriptions has been moved away from knowledge and thinking skills to give more emphasis to attitudes. The second four Key Competencies are asking for assessments that take into account not only the capacity of students to think about things and socio-cultural situations, but also the way they feel about, approach and interact with things and socio-cultural situations. There is an emphasis on attitudes in the last four Key Competencies, and taking attitudes explicitly into account in this trial is based on the belief that teachers know things about student attitudes that are not recorded in subject grades, and that potential employers are interested in teachers' assessments of both students' thinking skills and students' attitudes.

These considerations are related to the growing interest in Australia and internationally in developing what are called Enterprise Skills. The various conceptions of Enterprise Skills that are current overlap to a significant extent with the Key Competencies, but Enterprise Skills contrast with the Mayer conceptualisation of the Key Competencies in that Enterprise Skills give explicit emphasis to attitudes. The adaptation of the Key Competencies developed for this project has moved the Key Competencies to align them with the current interest in Enterprise Education.

Problems in assessing attitudes

It might be argued that only thinking skills can be assessed 'objectively' and that such skills are the only proper grounds for the assessment of students. In response it might be noted that all subject assessments are in some sense subjective, but it is an aim of this trial to see whether teachers can comfortably and reliably assess the mixture of skills and attitudes that are part of a comprehensive version of the Key Competencies.

The assessment of attitudes is, of course, a delicate matter. Attitudes cannot be readily encapsulated in a specific assessment task or specific pieces of behaviour, and assessing attitudes is always in danger of, and can easily degenerate into, a kind of tyrannous expectation of or demand for personal and ideological conformity. The kind of assessment of attitudes envisaged here is in some respects an assessment of personal characteristics and personality, but it is not an assessment of social conformity or ideology. Teachers are not to assess attitudes on the basis of merely personal views or preferences, but rather they are required to assess the functionality and/or effectiveness of students' attitudes. Making such assessments is a challenge to the professionalism of teachers (that is to their ability to be balanced, fair and 'objective'), but it is a challenge that the community, particularly employers, seem to want teachers to accept.

6 The nature of a Key Competencies assessment

This trial assessment might be described as based on contextualised and global, impression judgements made by individual teachers that are synthesised into overall school assessments.

The aim of this assessment is to produce overall school assessments recorded on an adaptation for schools of the three levels nominated by the Mayer Committee (see the **KC** Facets and Assessment Schema and the KC Facet Descriptors Overview). There are both principled and a practical reasons for aiming at overall school assessments.

If a Key Competencies assessment is to claim to be general, it must take into account as much of the school activity of a student as possible. And in practical terms, an overall judgement that arises from a range of teacher perspectives can be seen as escaping the limitations of any individual teachers, and, as a result, increasing the reliability of the assessment.

The judgements made by individual teachers in this trial are 'contextualised' in that they are made, primarily, on the basis of what teachers see in the course of teaching and assessing their normal courses. Teachers in this trial are not expected to teach differently or to undertake any new activities. They are expected to familiarise themselves with the Key Competencies, and to use them and the KC Facets and Assessment Schema and the KC Facet Descriptors Overview as the basis for global judgements towards the end of their courses. And that is all they are asked to do.

The judgements teachers are asked to make might be described as 'global impressions' because they are not based on specific tasks or instances of behaviour. They are to be overall and synthesised impressions that can take in as much evidence as the teacher can register. The assumption behind this assessment is that teachers do gain, to varying degrees, insight into the Key Competencies achievements of their students, and this assessment aims to draw on and formalise these impressions.

7 The Generic nature of Key Competencies judgements

It is crucial that the Key Competency judgements made by individual subject teachers in this trial are general rather than subject specific. Teachers will, of course, gain most of the knowledge of students from subject classes, but teachers in this trial are not being asked in assessing the Key Competencies to make judgements about that which is specific to their subject area. Key Competency judgements are not about students' performance in particular subject areas, they are judgements about what the particular teacher takes to be general about the abilities or the performance of a student. These judgements are broad or global inferences about what is taken to be typical of the student in most subject areas, and other activities.

Teachers of different subject areas have different perspectives on students and differences in their Key Competencies judgements are to be expected. But such differences should not result from teachers thinking they are assessing different things or because they assume that what is true of their subject or their personal interaction with the student is typical or general.

For instance, a student who has a gift for music may be a very positive member of an orchestra, but this does not mean that the music teacher who organises the orchestra should automatically assume or claim that such activity amounts to a high level of Working with others and in teams in general. The performance of such a student is the basis for claiming a level 3 Key Competency if the music teacher considers it a general characteristic of the student. It is, at least theoretically, possible that although the music teacher observes that student working well with others in the orchestra, she/he is doubtful about the student's ability to work well in teams in general, and so does not believe the student is typically at Level 3 in terms of team work. Other examples of the difference between subject-specific performances that might be recognised in subject grades or specific written comments, and the typical or general performances that are assessed in the Key Competencies levels might easily be elaborated. Performances in specific areas can and should be recognised in other forms of Key Competencies reporting, but the Key Competencies levels at issue in this project are not concerned with particular performances in specific areas.

The crucial point is that Key Competencies judgements are about what the teacher infers from what she/he observes about the general rather than the subject specific. This point is crucial not only for theoretical reasons. If an overall Key Competencies judgement is to be practical, valid and reliable, different teachers should think they are aiming to do the same thing, that is, aiming to assess what is typical and general about the performance of students. If different teachers produce many cases of inconsistent and irreconcilable judgements about individual students, then the notion of an overall Key Competencies judgement is invalid, as well as being unreliable and impractical.

Whether teachers can produce individual Key Competencies judgements that have a reasonable degree of consistency with those of other teachers is a crucial question to be explored by this trial.

8 The scope of the Key Competencies assessment

It is presumed that contact in subject classes is the basis of the Key Competencies assessment in this trial, but other information gathered from co- and extra-curricular activities and work placements can be taken into account. If such information is beyond what might be normally expected from a teacher in a particular subject area, note should be made (using the procedure described below) of the particular basis for the assessment.

For instance, if an art teacher had a particular view of the practical mathematical ability of a student gathered from working on the school play, or some other less expected observation, the art teacher would be wise to note the basis of the view if it is to be of much weight in comparison with a mathematics teachers in determining an overall level for this student on **Using mathematical ideas and techniques.**

9 Key Competencies Levels and the portfolio approach

The kind of assessment and reporting at issue in this project is very specific. The project is concerned with developing and reporting overall school assessments on the three levels described in the Mayer Report for assessing the Key Competencies.

This specific focus in not taken because it is presumed that assessment is the most valuable potentiality offered by the Key Competencies, or because reporting on three levels is the

only or the best way of reporting a Key Competencies assessment. The focus of this project is on reporting on three levels in the belief that this is a fundamental challenge for a widespread and even national assessment of Key Competencies.

The assessment considered in this project is compatible with much more comprehensive approaches to assessment, like the portfolio. Assessment on the 3 levels may well make a useful part of a portfolio assessment of Key Competencies and the portfolio would offer substance and exemplification of the overall judgements recorded in the KC levels. (This approach is recommended in the National Industry and Education Forum project on the portfolio assessment of Key Competencies.)

10 Making teacher judgements of Key Competencies in this trial

Participants in this trial are asked to gain a basic familiarity with the Key Competencies, the KC Facets and Assessment Schema and the KC Facet Descriptors Overview early in the process. In the course of their teaching and assessing during the second half of 1996 they are asked to give some thought to the extent to which and the ways in which the trial students are demonstrating Key Competencies. They might make notes during the course if they are so inclined, but such records are not required, and it is expected that many teachers can confidently make informed judgements about some of the Key Competencies for some students without recourse to formal reflection during the course or formal records.

Towards the end of the course teachers are asked to reflect formally on and record their judgements of the trial students on the KC Record Sheets. These records will offer the opportunity to make judgements on as many of the Key Competencies as the teacher wishes for each student. The judgements will be recorded on the following scale.

The KC Assessment Level Codes

Not Yet at Level One	NY
Not Yet/Basic	NY/B
Level One *Basic Achievement*	B
Basic/Medium	B/M
Level Two *Medium Achievement*	M
Medium/High	M/H
Level Three *High Achievement*	H
High Plus	H+

Teachers will use the KC Facet and Assessment Schema as a guide to determining a particular level out of the 8 codes on the KC Assessment Scale for a student. Teachers may offer no assessment of a KC for a student if they feel they have no insight into a particular Key Competency or a particular student.

Teachers may also record two additional pieces of information for the assistance of the person making the initial overall judgement. If the teacher imagines that her/his perspective might be seen as of less weight on a particular Key Competency or on the set because of the subject area in which she/he teaches, they may asterisk a particular assessment or group of assessments and make a note in the lines below the codes on particular issues to be taken into account in weighting that assessment in the overall decision. The example of the art teacher above would be in this category. Teachers might also asterisk particular assessments if they have particular reasons or pieces of evidence that would be usefully referred to the Overall Assessor. These notes may be the basis of the kind of school descriptive statement envisaged in the National Industry Education Forum portfolio assessment model.

Teachers may wish to record a judgement on a Key Competency for a student, but the addition of a question mark can be used to indicate a degree of uncertainty or a reservation on the part of the teacher making the assessment. Such a reservation might be explained in the space for notes below the codes.

The KC Facets and Standards Schema and the KC Facet Descriptors Overview indicate the basis on which individual judgements are to be made.

11 The Key Competencies Assessment Levels Framework for General Education

Two documents are offered in this project to support the assessment of the KCs by participating teachers. The first is the KC Facets & Standards Schema and the second is the KC Facet Descriptors Overview.

The KC Facets & Standards Schema

This one page document attempts to give the broadest of overviews of the basis for this assessment and the three levels for each KC. It is expected that this document will be the general framework on which participating teachers will make their assessments of KC level codes. The scope of each KC is further elaborated in the Facet Descriptors Overview.

The KC Facet Descriptors Overview

The statements in this document elaborate different aspects of each Key Competency.

It is intended that one descriptor and a level code will be included on the certificate of students to typify the main facet or characteristic of the student's performance in that competency.

The plan is that individual teachers will review and select one descriptor and level as the most typical of the student, and the Overall Assessor will look for agreement between teachers on the characteristics and the levels chosen for individuals by different teachers and will determine which Facet Description will be produced on the report form.

12 The Challenge of Key Competency or Generic Judgements: the Basis of the KC Level Judgements

The assessments made in this trial are within the framework of the KC Facets & Assessment Schema and the KC Facet Descriptors Overview, but these documents do not claim to explicitly delineate the different standards. These documents specify what issues are to be taken into account in the assessment, and macro levels of performance, but the limitations of these documents as definitions of standards are acknowledged. Teachers making KC assessments in this trial are involved in trying to envisage what might reasonably be expected from students in Year 11 and 12 programs, with the help of the standards schema and the facet descriptions.

The assumption behind these documents is that the judgements made in this project are stage related in that they are made explicitly about students in Grades 11 and 12 programs of general education. The judgements are made on the basis of participating teachers' knowledge and experience of students at this level and they are based on what teachers know and expect of students at this stage. The KC Assessment Facets Overview and KC Facets & Assessment Schema are offered as guidelines within which teachers are to exercise their judgements about students at this level.

It is proposed that assessments be reported in this project as judgements about the performance of students in general education programs. The basis of these judgements is that some students stand out, either because they exceed or do not meet usual and reasonable expectations. Students who do not meet basic expectations are judged to be below level one, and those who clearly distinguish themselves as high-level performers are at level three. These suggestions imply that most students will fall in levels one and two on a particular competency because they meet basic expectations but are not clearly distinguished. The distinction between Levels 1 and 2 is the difference between a basic level and some clear strength in a particular competency.

These presumptions can be expressed in another way to explain the process a teacher might use in making these Key Competencies judgements.

1 If a teacher is tempted to think a student is particularly strong or particularly weak in a Key Competency, they are faced with a decision between Levels 3 and 2 in the first example, and Not Yet Level 1 and Level 1 in the second example.

2 If a student is not clearly distinguished then the teacher is deciding between Levels 1 and 2.

3 In deciding between levels 1 and 2, teachers are looking for signs of strength that raise candidates above the basic expectation.

13 A process for making the overall school assessment

All the individual teacher judgements will eventually be reviewed by an Overall Assessor who will make an initial overall school assessment on the basis of the judgements of the various subject teachers.

The overall judgements will be made with the assistance of a computer program which will bring different assessments together and will make some adjustments if individual teachers are shown to be consistently harsher or more lenient than their colleagues. The program will also offer the Overall Assessor a suggested overall assessment or an indication that agreement is insufficient and reconciliation is needed before a result can be suggested.

The decision making process of the Overall Assessor will take into account the suggestions of the program, but the decision will not be an algorithm, like an average, and it will take into account the perspectives of different teachers in respect to a particular KC and any asterisked comments or question marks the teachers have recorded for the consideration of the Overall Assessor.

The aim of the Overall Assessor will be to determine as many overall results as possible without requiring any further information or discussion. If the range of levels offered for an individual student by different initial assessments is within three points on the eight point scale, then the Overall Assessor can comfortably determine and overall result. If the range of initial assessments is more than three points on the scale, careful consideration will have to be given to the range of levels, and a process of reconciliation may have to take place before an overall result can be determined.

Reconciliation of discrepant groups of assessments will take place through consultation between individual teachers and the Overall Assessor in cases where a particular teacher seems to be at odds with other teachers. Some reconciliations may be conveniently managed as a result of such one to one discussion. Where there is significant diversity between all teachers involved, reconciliation may be attempted through group discussion, or it may be decided that an overall report cannot be made efficiently.

When they are made, the overall assessments will be available for all teachers to review. Some of the assessments from different teachers will be irreconcilable, and the overall assessor may decide not to record a result or a set of results in these irreconcilable cases.

A meeting of interested parties may be convened to consider difficult cases so as to see if a reconciliation can be made between the different perspectives. Teachers may also raise queries about particular overall assessments at such a meeting. If a consensus or a majority view cannot be reached about the appropriate level or range for a particular student, a result will not be recorded for that student.

This sketch of the process of determining an overall assessment and reviewing it will be explained in greater detail later in the process.

The Key Competencies & Facets Overview

Each of the Key Competencies is a broad notion which has different aspects or facets. The Facets of the Key Competencies are nominated to give a manageable working definition of the broad notion.

Although the aim of this assessment is to report a single level for each Key Competency, it is intended that the additional reporting of a particular facet of the Key Competency will carry some added information for readers of the report. The aim of this material is to identify different aspects of each Key Competency as an aid to understanding and assessing them. It is also intended that the Facets will be the basis for more qualitative reporting of the reasons for awarding students overall levels.

What is meant by a Facet of a KC?

The aim of these three statements about each KC is to:

- map the Key Competency and indicate the issues to be taken into account in assessing it; and
- give a means of reporting the aspect of performance that is typical of the student.

The facets are intended to be:

- part of a general but coherent notion;
- a different aspect from the others in the group; and
- a key distinguishing feature of performance on the Key Competency.

Guidelines for the Key Competencies Assessment

Participants are asked to review the facets and give a Facet level on a 4 point range for each Facet of each KC. Participants are also asked to determine an on-balance and overall level for each KC on the 8 point range. In deciding the overall level, participants are also to select and record a Facet descriptor code that is a particular strength or best typifies the performance of the student on a competency. This process means the KC Level is an on-balance judgement based on an assessment of each of the facets of the KC.

Step 1 Facet Level Codes

The first step is to assess each Facet using the 4 point range of **Not Yet** (NY), **Basic** (B), **Medium** (M) and **High** (H).

Step 2 Overall Level Codes

The second step is to determine an overall level on the 8 point range of **Not Yet Basic** (NY), **Not Yet Basic/Basic** (NY/B), **Basic** (B), **Basic/Medium** (B/M), **Medium** (M), **Medium/High** (M/H), **High** (H), **High Plus** (H+).

There may be differences in the performance of students on different Facets. These differences are to be balanced against one another in determining an overall level.

The Three Levels of Performance

- Level 3 and Below Level 1 are used when students clearly exceed, or do not reach, usual and reasonable expectations.

- Level 2 is for those students who show definite strength in a KC but do not clearly exceed expectations.

Step 3 KC Commentary Codes

The Commentary Codes are ways of adding further information for the Overall Assessor to consider in making an overall decision.

One or a number of questions marks or ticks (??, ✓✓) can be used to indicate degrees of uncertainty or certainty about the assessment offered.

Arrows up and down next to an assessment (⇑ ⇓) indicate an inclination to vary the level in the direction of the arrow.

Asterisks (**) can be used to signal the inclusion of a comment explaining the level judgement or commenting on the performance of the student in a particular area.

Step 4 Typical Facet Codes

This is the code (A,B,C) for the one Facet of each KC that is a **particular strength or best typifies** a student, and was most significant in the level decision. It is recorded next to the level code.

Each of the Key Competencies has a number of facets which can be seen as the separate criteria that map the scope of the KC.

The aim of the assessment is to record an overall position on the range from **NY** to **H+**. The overall decision is an on-balance judgement taking into account the student's achievements on each facet. Participants are also asked to record a judgement on the range from **Not Yet** to **High** for each of the Facets in making an overall on-balance judgement.

A second part of the assessment is to record the facet of the student's achievement that is strongest or most typical. The strongest or most typical facet of a student's achievements will be reported as well as a level to add further information.

KC1	**Facets of Collecting, Analysing & Organising Ideas & Information**								
1.A	comprehending and interpreting ideas and information		Not Yet		Basic		Med		High
1.B	analysing and evaluating ideas and information		Not Yet		Basic		Med		High
1.C	synthesising and developing ideas and information		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facet A focuses on the more receptive acts of making sense of ideas and information. It envisages a strength in understanding ideas and information.

Facet B focuses on critique and assessment. It envisages a strength in critical thinking.

Facet C focuses on the manipulation of ideas and information. It envisages a strength in elaborating or applying information.

KC2 Facets of Communicating Ideas & Information

The facets of this Key Competency are separated into spoken and written communication.

KC2.a	**Oral communication**								
2.a.A	speaking in less formal situations		Not Yet		Basic		Med		High
2.a.B	speaking in more formal situations		Not Yet		Basic		Med		High
2.a.C	using speech to interact with other people		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facet 2.a.A and 2.a.B are separated into less formal and more formal uses of oral communication.

Facet 2.a.C envisages a strength in using oral communication to establish a rapport with and respond to other people.

KC2.b	**Written communication**								
2.b.A	writing clearly and accurately		Not Yet		Basic		Med		High
2.b.B	writing confidently and fluently		Not Yet		Basic		Med		High
2.b.C	using writing to explore ideas and information		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facet 2.b.A envisages a strength in the correct and precise expression of ideas in writing.

Facet 2.b.B envisages a strength in using writing with flair and expressiveness.

Facet 2.b.C envisages a strength in using writing to think and develop ideas.

KC3	**Facets of Using Mathematical Ideas & Techniques**								
3.A	using mathematical techniques to represent and analyse ideas and information		Not Yet		Basic		Med		High
3.B	understanding mathematical principles and procedures		Not Yet		Basic		Med		High
3.C	applying mathematical knowledge to problems		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facet A is concerned with the use of mathematical methods for representing data and analysing issues.

Facet B is concerned with the fundamental ideas and techniques of numeracy and mathematics.

Facet C is concerned with the practical use of mathematical understanding.

KC4	Facets of Using Cultural Understandings								
4.A	comprehending and interpreting social issues		Not Yet		Basic		Med		High
4.B	developing a coherent position on social issues		Not Yet		Basic		Med		High
4.C	using cultural understanding to achieve goals		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	Y/B	B	B/M	M	M/H	H	H+

Facet A envisages a strength in understanding and making sense of socio-cultural ideas.

Facet B envisages a strength in evaluating socio-cultural ideas and developing views about them.

Facet C envisages a strength in applying cultural understanding to deal with situations or to solve problems.

KC5	Facets of Problem Solving								
5.A	shows focus and persistence		Not Yet		Basic		Med		High
5.B	shows independence and responsibility		Not Yet		Basic		Med		High
5.C	shows initiative and creativity		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

These Facets identify different attitudes or personal characteristics that play a part in problem solving. They emphasise the way students approach or deal with challenges or problems.

KC6	Facets of Using Technology								
6.A	seeks technological experiences and challenges		Not Yet		Basic		Med		High
6.B	is thorough and persistent in dealing with technology		Not Yet		Basic		Med		High
6.C	can understand technological systems		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facets A envisages a positive attitude to technology.

Facet B envisages an approach to using technology.

Facet C deals with conceptual understanding of technology.

KC7	Facets of Planning and Organising Activities								
7.A	identifying goals, priorities and strategies		Not Yet		Basic		Med		High
7.B	implementing plans and strategies		Not Yet		Basic		Med		High
7.C	seeks planning and organisational experiences and challenges		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

These facets are personal characteristics that indicate different strengths in dealing with organisational challenges.

Facet A envisages a strength in planning activities.

Facet B envisages a strength in organising and executing activities.

Facet C envisages a positive attitude to planning and organisational challenges.

KC8	Facets of Working with Others and in Teams								
8.A	adapts to group expectations		Not Yet		Basic		Med		High
8.B	leads or facilitates group processes		Not Yet		Basic		Med		High
8.C	assists and supports others		Not Yet		Basic		Med		High
	Overall on-balance judgement	NY	NY/B	B	B/M	M	M/H	H	H+

Facet A is the ability to adjust to and fit in with a group.

Facets B and C are about the way people work in and contribute to groups.

Appendix 6
The SBKCLAP Overall Assessment Algorithm

The table below presents the decisions for varying numbers of teachers (column 1), the range of scores initially given by the teachers (column 2), and the bottom and top points of a reporting range according to the algorithm. The same range of scores and outcomes would be applied at any point in the actual score range (e.g. 2-4 or 6-8).

Number of Teachers	Range of Scores	Report bottom	Report Top
1	1	1	1
2	11	1	1
2	12	1	2
2	13	1	3
2	14	No	Result
3	111	1	1
3	112	1	2
3	122	1	2
3	113	1	2
3	123	1	3
3	133	2	3
3	114	1	3
3	124	1	3
3	134	2	4
3	144	2	4
4	1111	1	1
4	1112	1	1
4	1122	1	2
4	1222	2	2
4	1113	1	2
4	1123	1	3
4	1223	1	3
4	1133	1	3
4	1233	2	3
4	1333	2	3
4	1114	1	2
4	1124	1	3
4	1224	1	3
4	1134	1	3
4	1234	2	3
4	1334	2	4
4	1144	No	Result
4	1244	2	4
4	1344	2	4
4	1444	3	4
5	11111	1	1
5	11112	1	1
5	11122	1	2
5	11222	1	2

5	12222	2	2
5	11113	1	1
5	11123	1	2
5	11223	1	3
5	12223	2	2
5	11133	1	3
5	11233	1	3
5	12233	2	3
5	11333	2	3
5	12333	2	3
5	13333	3	3
5	11114	1	1
5	11124	1	2
5	11224	1	3
5	12224	1	3
5	11134	1	3
5	11234	1	3
5	12234	1	3
5	11334	1	3
5	12334	2	4
5	13334	2	4
5	11144	1	3
5	11244	1	3
5	12244	1	3
5	11344	2	4
5	12344	2	4
5	13344	3	4
5	11444	3	4
5	12444	3	4
5	13444	3	4
5	14444	4	4
6	111111	1	1
6	111112	1	1
6	111122	1	1
6	111222	1	2
6	112222	2	2
6	122222	2	2
6	111113	1	1
6	111123	1	1
6	111223	1	2
6	112223	2	2
6	122223	2	2
6	111133	1	1
6	111233	1	3
6	112233	1	3
6	122233	2	3
6	111333	1	3
6	112333	1	3
6	122333	2	3
6	113333	3	3
6	123333	3	3

6	133333	3	3
6	111114	1	1
6	111124	1	1
6	111224	1	2
6	112224	1	2
6	122224	2	2
6	111134	1	1
6	111234	1	3
6	112234	1	3
6	122234	1	3
6	111334	1	3
6	112334	2	4
6	122334	2	3
6	113334	1	3
6	123334	2	3
6	133334	3	3
6	111144	1	1
6	111244	1	3
6	112244	1	3
6	122244	2	4
6	111344	1	3
6	112344	2	3
6	122344	2	4
6	113344	2	4
6	123344	2	4
6	133344	3	4
6	111444	No	Result
6	112444	2	4
6	122444	2	4
6	113444	2	4
6	123444	2	4
6	133444	3	4
6	114444	4	4
6	124444	4	4
6	134444	4	4
6	144444	4	4

Appendix 7 The School-based Key Competencies Levels Assessment Project Final Report

THE SCHOOL BASED

KEY COMPETENCIES

LEVELS ASSESSMENT PROJECT

FINAL REPORT

A project researching the school based assessment of the Key Competencies

funded by the Department of
Employment, Education, Training and Youth Affairs

Doug McCurry
&
Jennifer Bryce

April 1997

This project was supported by a grant under the Key Competencies Program of the Commonwealth Department of Employment, Education, Training and Youth Affairs. The views expressed here are those of the authors and do not necessarily represent the views of the Department.

CONTENTS

APPENDICES

The following appendices are included at the end of the report.

Acknowledgments

The primary authors of this report, Doug McCurry and Jennifer Bryce, are pleased to acknowledge the work of Margaret Wu and Greg Macaskill of ACER in analysing the assessment data and in developing the computer program for assessing and reporting the Key Competencies. We are also pleased to acknowledge the work of Dr. Chris Reynolds to produce the report in Appendix 10 - *Reactions of employers to a sample report based on the Mayer Key Competencies.*

SUMMARY

This project undertaken by ACER for the Department of Employment, Education, Training and Youth Affairs aimed to explore the school based assessment of levels of student performance on the constructs nominated by the Mayer Committee in their report: **The Key Competencies** (1992).

In a trial assessment 110 teachers in 10 schools across the country were asked to give global assessments of each of the Key Competencies for a total of 629 Year 11 students they were currently teaching using the Key Competencies framework developed for the purpose during the project. The teachers and the students were not expected to undertake any new or different work, and the assessment was made on the basis of what teachers learned about their students from their usual work programs. The project aimed to produce overall levels of performance for individual students from a range of different teachers' assessments.

The project also surveyed the reactions of employers to a report format including the summary levels of performance developed in the assessment trial.

At the end of the trial assessment, participating teachers were interviewed, and they were also asked to complete a written survey (see Part 3). This gathering of data from the participant teachers showed that:

1. the Key Competencies, as presented in the project, were generally seen as a satisfactory conception of employment-related competencies, although there were some concerns about the intelligibility and appropriateness of the Using Mathematical Ideas and Techniques, Using Technology and the Using Cultural Understandings Key Competencies;

2. the facets of the Key Competencies developed by the project were generally seen as defining coherent concepts, and participants generally found the assessment framework developed by the project clear and intelligible;

3. participating teachers were able to make global impression judgements about Key Competencies performance on the basis of their knowledge of students without undertaking new or different tasks;

4. teachers were able to use the assessment procedures developed by the project, and most found the task easier than they had expected; and

5. there was significant support among participants for the view that a Key Competencies assessment of the kind trialled in the project could be both cost effective and report useful information.

The results of the teacher assessments of students were compared and analysed (see Part 4). The analysis of the assessment data from the trial showed that:

6. participating teachers were able to discriminate among their students and spread them across the mark range using the Key Competencies framework;

7. significant numbers of teachers did not assess Using Mathematical Ideas and Techniques, Using Technology and Using Cultural Understandings, but almost all participants were able to assess the other Key Competencies;

8. there were some systematic variations in the way the teachers used the mark range, (there were some comparatively harsh and lenient teachers), and there were differences

in the consistency of individual teachers' assessments with their colleagues (there were some relatively consistent and some inconsistent teachers);

9. there was quite a high level of agreement between different teachers assessing the same student on the whole range of Key Competencies (a level of agreement comparable to that attained in external assessments of single pieces of work on similar mark ranges);

10. the level of agreement between teachers about individual students in this trial makes it realistic to construct overall and cross-curricular assessments of Key Competencies;

11. the data offered little indication that the perspectives of teachers in different subject groupings differ substantially or systematically from teachers in other subject groupings;

12. the data offered little indication that the perspectives of female and male teachers differ substantially or systematically in the way they assess either girls or boys; and

13. there was a fairly consistent pattern in the way individual schools assessed their students in comparison with other schools.

The analysis of the written and telephone survey of employer reactions to a trial Key Competencies report (see Part 5 and Appendix 10) showed that:

14. respondents found the Key Competencies to be a good basis for schools reporting student performance to employers, but they require other information as well;

15. the information provided in the Sample Report was good, but for most its presentation was too complex and unclear;

16. levels of performance information was welcomed by almost all respondents;

17. the concept of facets and the level descriptors provided was widely welcomed, and the concept of distinctive comments in a school statement was strongly supported;

18. the monitoring of assessments produced by schools was seen as highly desirable by the great majority of respondents;

19. there is clear support for the inclusion of a Student Statement but concerns were voiced about how it could be authenticated and therefore equitably regarded;

20. there was general support for the inclusion of Specific Skills in the list provided by the student; and

21. there was general support for the inclusion of a list of portfolio items, especially if the selection of them is targeted to a particular job application and related to the Key Competencies.

The results of this trial suggest that the Key Competencies can be seen as generalised notions that may hold their meaning across different subject areas or domains, and that they can be operationalised in a coherent fashion. The trial has shown that teachers within a school can discriminate among the Key Competencies performances of their students, and that the assessments of different teachers can be meaningfully and usefully synthesised into overall levels that can be reported as whole school judgements.

School Based Key Competencies Levels Assessment Project Overview

This project aimed to explore the assessment and reporting of the Key Competencies nominated by the Mayer Committee. The assessment trial component of the project was designed to explore the assumptions of the Mayer Committee that meaningful judgements could be made about the Key Competencies performance of students by individual teachers, and that these assessments could be synthesised and reported as meaningful summaries of the Key Competencies achievements of students.

In particular, the project aimed to see whether teachers in general education programs could make Key Competency judgements without undertaking new or different work, and whether there is sufficient agreement between different teachers to make it practical to produce an overall assessment from a provider for a student.

This report gives an overview of this project and presents the qualitative and quantitative results of the assessment trial. It also presents the results of the second component of the project, a questionnaire and telephone survey soliciting employer reactions to a comprehensive portfolio based report format developed from the work of the Business, Industry, Key Competencies & Portfolios Project of the National Industry Education Forum (1996).

The Key Competencies Controversy

Since the publication of the report of the Mayer Committee a number of reservations and doubts have been expressed about the assumptions made and the proposals offered by the Mayer Committee.

At a theoretical level it has been argued that Key Competencies 'do not exist'. The sophisticated versions of this argument have drawn on a body of research that has questioned the notion of generic skills and, particularly in the research comparing novices and experts, has sought to show that competence is domain or context specific. The implication of these arguments is that it is not meaningful or useful to hypothesise or attempt to assess generic skills because the performance of individuals will differ markedly from domain to domain in a way that makes generalisation impossible.

This project has set out to test these issues by exploring whether teachers can make judgements that are not domain specific, on the assumption that congruence or agreement between teachers of different subjects offers evidence about the extent to which students' performances can be usefully seen as generic and can be generalised.

At a more practical level, it has been argued by critics of the Key Competencies that teachers will not be able to make generic Key Competencies judgements, or that they will not be able to make them reliably without a new and time-consuming assessment regime.

So, part of the argument that questions the feasibility of Key Competencies assessment is based on the objection to the very notion of generic skills, and another part of this argument is based on the view that Key Competencies assessment is not practical or cost effective.

This project attempted to explore these issues by designing a Key Competencies assessment regime and implementing it. It indirectly addresses the theoretical reservations about the assessment of the Key Competencies by attempting to see whether teachers find the process of making Key Competencies judgements makes sense in practice, and whether results of Key Competencies assessments are consistent. And the project directly attempts to test whether a practical and cost effective Key Competencies assessment can be designed and implemented.

The aims of the School Based Key Competencies Levels Assessment Trial

The thinking behind the project can be seen in the report written for The Department of Employment, Education, Trade and Youth Affairs by McCurry (a relevant excerpt is presented in Appendix 1) about the assessment of Key Competencies, and this document shaped the Project Brief which is presented in Appendix 2.

In this trial 110 teachers in 10 schools around the country were asked to give a global assessment of 629 Year 11 students they were currently teaching using the Key Competencies framework developed for the purpose during the project. The teachers and the students were not expected to undertake any new or different work, and the assessment was made on the basis of what teachers learned about their students from current work programs.

Participants in the trial were asked to develop an understanding of the Key Competencies, and towards the end of the year they were asked to make a decision about which of eight levels they saw the trial students as reaching on each Key Competency. The assessments of the different teachers in each school were synthesised into an overall assessment for a student using an algorithm, and these results were reviewed by a member of the school staff acting in the role of an Overall Assessor.

The object of this work was to design and trial a reliable and cost effective method for producing overall assessments on eight levels of performance for each of the Key Competencies. These assessments were presented to participants as particularly intended to assist students in transition from school to work.

The aims of the trial were to see if:

1. the participating teachers could make global Key Competencies judgements comfortably;
2. different teachers were reasonably consistent in the way they judged the performance of students on the Key Competencies levels; and
3. the procedures for making overall assessments were practical and cost effective.

1 Part 2 The SBKCLAP Research Procedure

SELECTION OF SCHOOLS

The scheme for making assessments of Key Competencies by teachers' global impression judgements developed by this project was trialled in 10 schools representing government and non government systems and different state systems of education. As it was thought uneconomic to select schools from all Australian states, the final selection comprised schools from New South Wales, Queensland, South Australia and Victoria.

STATE	TYPE OF SCHOOL	NUMBER
New South Wales	government	2
Queensland	government	3
South Australia	catholic	3
Victoria	government	2
	TOTAL	**10**

Appropriate authorities for independent, government and catholic schools were approached and asked to recommend schools that would meet the following criteria:

- not involved in other projects concerning the Mayer Key Competencies;
- with above average numbers of Year 11/12 students moving from school to work rather than to tertiary education;
- likely to be interested and willing to take part in the assessment activity;
- large enough to have at least 80-100 students in Year 11.

It seemed desirable to have schools not necessarily familiar with the Mayer Key Competencies as this would test the clarity of the project materials. Given the employment-related nature of the Key Competencies, it was thought likely that the project would have greater appeal for schools with large numbers of students aiming to find work immediately after Year 11 or 12 rather than those schools preparing the bulk of their school leavers for tertiary studies. Year 11 was targeted to satisfy an interest in assessing students' Key Competencies close to the point of leaving school whilst not imposing on the sometimes onerous time-table of the last part of Year 12. The original target was to have assessments for about 700 students (80 students in 9 schools) with each student assessed by at least four different teachers. The final number of students was 629.

Each authority operated in a slightly different way, but as a result of this request authorities provided names of:

- 6 government schools in Victoria
- 4 catholic schools in South Australia
- 3 government schools in New South Wales.

Three government schools in Victoria were selected, and all other schools suggested were contacted. This involved a letter to the principal outlining the project, mentioning that money would be paid for teacher release and outlining other ways in which it was believed a school would benefit from participating, viz.:

- the school community would become familiar with the Key Competencies;
- teachers would become familiar with the Key Competencies and would receive inservice in the assessment procedures;
- teachers would become aware of the cross-curricula competencies developed in their students.

Excellent co-operation was received from authorities, but schools were being asked to release senior teachers in the latter, most stressful part of the year. Indeed, the trial was described as 'coldstart and minimal' because the schools would know comparatively little about the Key Competencies and they would be given comparatively little support in doing the assessment. As would be expected in these circumstances, the selection of schools did not run exactly to plan. One of the Victorian schools selected was already involved in a Key Competencies project, so preferred to decline to take part in this one. A second school undertook the initial inservice then had to discontinue due to internal problems. Only one of the three government schools in New South Wales was able to participate. This was mainly because of industrial issues at the time. Typical of the response from New South Wales was the text of this facsimile:

> *We feel that we cannot at this stage participate given the short notice to consult with staff teaching Year 11 and the current industrial issues in New South Wales. Personally I feel the project would be of great value for professional development in Assessment and Reporting.*

Three of the four catholic schools in South Australia were able to take part. This left one school in New South Wales, one in Victoria and three in South Australia - only five schools, and it was not economically feasible to travel to New South Wales for one school. Another school in Victoria was prepared to take part at short notice. It was decided to approach Queensland government schools instead of going to New South Wales. After permission was obtained to go to three Queensland government schools, one New South Wales high school (not previously approached) expressed interest in being involved in the project. It was decided that since New South Wales could be visited whilst en route to Queensland it would now be worth including two New South Wales schools - the one that initially gave permission and the one that had recently expressed interest. So, although the original intention was to have only nine schools in the trial, the final count was ten.

It is important to stress that although schools were selected to meet the requirements outlined above, teachers were not specially selected. Indeed, some may have been reluctant to take part. Some knew very little of the Key Competencies prior to the trial and, as can be seen in Part 3, they were not necessarily advocates of the Key Competencies.

The time-table for these school involvements was:

TIME OF YEAR	KIND OF INVOLVEMENT	PURPOSE
July/August	researchers to visit schools	☐ familiarise with Key Competencies ☐ familiarise with proposed assessment strategies ☐ confirm staff to be involved and give them some time to think about the tasks and confirm the appointment of 'overall assessor' select students to be assessed
September	negotiate with year co-ordinator or other appropriate staff member	notify of any changes to assessment framework and answer any queries before staff undertake the assessment
October/November	researchers to visit schools	☐ staff undertake the assessment ☐ overall assessor responsible for sending completed assessment sheets to the researchers
October/November		☐ distribute evaluation questionnaires and interview participants ☐ work with overall assessor on range of teachers' assessments
November/December	researchers to visit schools	

THE KEY COMPETENCY LEVELS ASSESSMENT ADVISORY GROUP

An Advisory Group for the project was identified from different groups involved in the Key Competencies program (the participants are outlined in Appendix 3), and the first meeting of the group was held during the inception phase in June 1996 and a second meeting of the advisory group was held in March 1997 to review the results of the trial and survey .

The aim of the first meeting was to review the framework developed for the assessment trial (see Appendix 4). The Advisory Group members were asked to put themselves in the position of the participating teachers so as to help the project team refine the documents and the assessment procedures before they were introduced to the participants.

The discussion of the framework during the Advisory Group meeting noted that the presumptions and assumptions on which this project is based are challenging to more conventional approaches to assessment, but it was recognised that these positions are

deliberately taken in the belief that they make Key Competencies assessment in general education conceptually coherent and practical.

There was particular concern expressed by a number of members of the Advisory Group that the trial depends on the belief that an assessment of Key Competencies is possible on the basis of current practice and without changes to current curriculum and pedagogy. This is, of course, a fundamental assumption of this trial, and while it is assumed that meaningful assessments can be derived from current programs, it is not based on the view that current programs will produce an optimal assessment of the Key Competencies. **This trial does not assume that the implementation of a Key Competencies assessment need not entail changes to current curriculum and pedagogy, but it does assume that Key Competencies assessment is not at odds with current practice, and that current practice gives a basis for a certain, limited kind of Key Competencies assessment.**

THE FIRST SCHOOL VISIT

Prior to the first school visit each teacher was sent 'Participant Background Notes' (Appendix 4). These outlined the areas to be covered in the first inservice session. It was not imperative that teachers read the notes before the session. They were provided because it was thought that some teachers might be interested to know about the project prior to the first visit.

The aims of the first school visit were to

- give a brief overview of the development of the Mayer Key Competencies including discussion of the climate in which they developed;
- outline propositions which underpin the approach to assessment;
- familiarise teachers with the definitions of the Key Competencies to be used in the project, including an introduction to the 'facets' used to map the Key Competencies;
- discuss the process of making global impression judgements; and
- outline the purpose and role of the overall assessor.

Assumptions underpinning the research

The participant briefing notes outline the views underpinning the approach to assessment taken in this project that the Key Competencies can be meaningfully and realistically assessed in general education programs, and that they can be assessed in a minimal but useful fashion without changes to current school programs and without placing substantial new burdens on teachers and students. The following propositions are presented in The Briefing Notes:

1. Although in some respects the Key Competencies are unfamiliar to teachers, they can make sense to teachers, and they can be assessed globally by teachers.

2. The Key Competencies are part of general education; teachers know things about the Key Competencies performance of their students; and the Key Competencies can be assessed as part of current school programs.

3. Key Competencies assessments are not in competition with or a replacement for subject-based assessments. They offer information that is in addition to that offered by subject-based assessments.

4. Key Competencies assessments are aimed particularly at offering information about students who are in transition to work, and they touch on skills and attitudes not always or readily included in subject assessments.

5. The most realistic way of reporting summary judgements of Key Competencies performance in general education is on the basis of global judgements made by individual teachers about the general skills and attitudes of the students they teach, with these individual judgements being subsequently synthesised into and reported as overall school levels.

6. Judgements about Key Competencies performance should be broad and general inferences about the way students are likely to perform in the future rather than judgements about students' performances in specific courses. Because they are general or generic, Key Competencies judgements can be synthesised into collective school judgements.

Adaptation of the Mayer Key Competencies

The Mayer Committee conception of the Key Competencies was the basis of this project, but the Key Competencies were re-ordered according to whether their conception was seen as more or less cognitive, and they were more precisely defined for the purposes of assessment by means of facet descriptors.

The notion of assessing attitudes was rejected by the Mayer committee, but in envisaging how competencies such as Planning and Organising Activities or Working in Teams are to be assessed it makes sense to see them in terms of kinds of approaches or attitudes rather than trying to assess specific performances.

Solving Problems had been criticised for its generality - but seeing it as an approach or attitude to solving problems makes its assessment more feasible. This is also the case for Using Technology, otherwise one asks - what technology? and becomes involved with issues of equity. On the other hand, Using Cultural Understandings, the most controversial of the Key Competencies had been so partly because some had seen it as dealing with socio-political value judgements. To circumvent this problem, it was interpreted in a more cognitive way in this project. These issues are discussed at greater length in the briefing notes (Appendix 4).

The principled reasoning for explicitly introducing attitudes into the Key Competencies assessment is to recognise that competence in employment is usually defined as involving knowledge, skills *and attitudes*. The briefing notes stress that the kind of assessment of attitudes envisaged for this project is an assessment of personal characteristics rather than an assessment of social conformity or ideology.

The resulting re-ordering of the Key Competencies is given below (changes to the original Mayer Committee wording are given in italics):

Collecting, Analysing and Organising Information

The capacity to locate information, sift and sort information in order to select what is required and present it in a useful way, and evaluate both the information itself and the sources and methods used to obtain it.

Communicating Ideas and Information in *Speech and Writing*

The capacity to communicate effectively with others *using speech and writing.*

Using Mathematical Ideas and Techniques

The capacity to use mathematical ideas, such as number and space, and techniques, such as estimation and approximation, for practical purposes.

Using Cultural Understandings

The capacity to understand and interpret socio-cultural ideas, and to reason plausibly about personal and inter-personal issues.

Solving Problems

The capacity *and the attitude* needed to apply problem-solving strategies in purposeful ways, both in situations where the problem and the desired solution are clearly evident and in situations requiring critical thinking and a creative approach to achieve an outcome.

Using Technology

The capacity *and the attitude* needed to apply technology, combining the physical and sensory skills needed to operate equipment with the understanding of scientific and technological principles needed to explore and adapt systems.

Planning and Organising Activities

The capacity *and the attitude* needed to plan and organise one's own work activities, including making good use of time and resources, sorting out priorities and monitoring one's own performance.

Working with Others and in Teams

The capacity *and the attitude* needed to interact effectively with other people both on a one-to-one basis and in groups, including understanding and responding to the needs of a client and working effectively as a member of a team to achieve a shared goal.

The briefing notes point out that the first four Key Competencies emphasise kinds of thinking skills - the same sorts of thinking skills that are the basis of the school curriculum, or, more precisely, the skills assessed in school grades. The second four can also be thought of as thinking skills, but the emphasis in the descriptions has been moved away from knowledge and thinking skills to give more emphasis to attitudes. The second four Key

Competencies are asking for assessments that take into account not only the capacity of students to think about things and socio-cultural situations, but also the way they feel about, approach and interact with things and socio-cultural situations.

The facet descriptors

In order to achieve a coherent notion of each Key Competency so that each teacher would be assessing the same aspects of a student's performance, it was necessary to define more clearly the Key Competencies. This was done by developing facet descriptors which aimed to identify aspects of the Key Competencies which were each different, yet a part of the coherent whole of that competency. The facets of each Key Competency thus aimed to:

- map the Key Competency and indicate the issues to be taken into account in assessing it;
- give a means of reporting the aspect of performance that is typical of the student.

The facets were intended to be:

- part of a general but coherent notion;
- a different aspect from the others in the group for that Key Competency;
- a key distinguishing feature of performance on the Key Competency.

The facets of each Key Competency are outlined in Appendix 4.

Assessment by making global impression judgements

Teachers were asked to make judgements about students that were not based on specific tasks or instances of behaviour but judgements that were overall and synthesised impressions taking in as much evidence as a teacher can register. The generic nature of the Key Competencies is stressed in the background notes. It is pointed out that teachers will gain most of their knowledge about students from subject classes, but that the judgements to be made are concerned with what a teacher takes to be general about the abilities or performance of a student - inferences about what is typical of the student in most subject areas and other activities. The judgements for this project concerned students in Years 11 and 12 general education programs. Teachers were told to base their judgements on their knowledge and experience of students at this level and on what they know and expect of students at this stage.

Using the assessment framework

The Mayer report proposes that Key Competencies be assessed over three levels of achievement. These levels are used in the scheme for this project: Level 1 (basic achievement at grades 11 and 12), Level 2 (medium achievement in grades 11 and 12) and Level 3 (high achievement in grades 11 and 12). To facilitate the assessment, level descriptors were written and the levels were broken into 'level codes' which enabled teachers to assess over an eight-point scale. The level codes are outlined below.

The KC Assessment Level Codes

Not Yet at Level One	NY
Not Yet/Basic	NY/B
Level One *Basic Achievement*	B
Basic/Medium	B/M
Level Two *Medium Achievement*	M
Medium/High	M/H
Level Three *High Achievement*	H
High Plus	H+

Teachers were encouraged to make their judgements by thinking of how some students stand out, either because they exceed or do not meet usual and reasonable expectations. Students who do not meet basic expectations are judged to be below level one, and those who clearly distinguish themselves as high level performers are at level three. So most students will fall in levels one and two on a particular competency because they meet basic expectations but are not clearly distinguished. The process was further outlined in the following terms:

1. If a teacher is tempted to think a student is particularly strong or particularly weak in a Key Competency, they are faced with a decision between Levels 3 and 2 in the first example, and Not Yet Level 1 and Level 1 in the second example.
2. If a student is not clearly distinguished then the teacher is deciding between Levels 1 and 2.
3. In deciding between Levels 1 and 2, teachers are looking for signs of strength that raise candidates above the basic expectation.

The role of overall assessor

On the first school visit, the role of overall assessor was outlined. In most cases an appropriate staff member had already been selected - usually the Year 11 co-ordinator who is familiar with staff and students at that level.

Initially it was planned that the overall assessor would review teachers' judgements with the aid of a computer program. For various reasons it was later decided to follow the same process using pen and paper and to leave the implementation into software until after the assessment phase of the project had been completed.

The main aim of the overall assessor in this scheme was to bring together the judgements of teachers into an assessment of a student which would range over no more than three points of the eight-point scale. This was to be done not by averaging, but by taking into account the

perspectives of different teachers and noting comments or other indications they had made. Sometimes discrepant assessments would need to be reconciled by discussion with teachers and it was recognised that sometimes reconciliation might be impossible, in which case no assessment would be given for that student on the particular Key Competency.

Summary: the first school visit

A lot of material was covered in the first school visit. This was done in a half day's inservice covering the points outlined above. In several schools teachers were quite keen to start using the assessment framework and they worked in pairs doing practice assessments for students they knew. A benefit of this school visit that had not been anticipated was that teachers from a variety of subject areas were brought together and spent several hours working together. In large schools teachers said that they more usually undertake inservice in subject groups and they appreciated the cross curricular exchange.

By the end of the first school visit it was expected that teachers would have a clearer conception of the Key Competencies and familiarity with the assessment scheme they would use at the end of the year.

SELECTION OF STUDENTS TO BE ASSESSED

The assumption that in a school with 80 - 100 Year 11 students it would be reasonably easy to find a group of 6 or 7 teachers who taught or knew through other activities most of these students was naive. Each student was to be assessed on the Key Competencies by at least four different teachers. In reality, setting this up was extremely complex because there are so many different units at Year 11 level taught by different teachers, so there were very few teachers who had taught or 'knew' a large body of the Year 11 students.

Because of this complexity some changes had to be made and some teachers had to be recruited to participate in the assessments even though they had not attended the initial inservice session. (Unintentionally, this provided an opportunity to test the clarity of the materials - even those teachers called in 'at the last minute' were able to come to terms with the procedure and undertake the assessments.)

In the end some cases where a minimum of three rather than four teachers assessed a student had to be included. Overall, 110 teachers assessed 629 students.

THE SECOND SCHOOL VISIT

The conception of the overall framework did not change from the time it was presented in the first school visit, but some modifications were made to the facet descriptors in the light of teachers' suggestions and reflections of the researchers. There were three facets for each Key Competency and it was decided to ask teachers to assess each facet (as well as indicate a student's strongest facet). This was for research purposes - to see, for example, whether one facet of a Key Competency tended to get higher gradings than the other two. It is not recommend that this be a part of the final assessment model.

The second school visit involved meeting with teachers and ensuring that they understood the assessment framework. In some schools this also involved meeting with the overall assessor and confirming which students were to be assessed.

Because of the unavoidably slow procedure in selecting schools (described above), there were three schools where the first and second visit had to be collapsed into one. This meant

that teachers did not have much time to think about their students in relation to Key Competencies, but there were no obvious difficulties at these schools.

Copies of the final versions of standards schema and facet descriptors are given in Appendix 4.

THE THIRD SCHOOL VISIT

Teachers were given a date in November by which time all assessments were to be completed. The overall assessor supervised the return of all completed assessment documents to ACER by that date. There was a short amount of time in which to process the assessment materials into a format that could be checked by the overall assessor - a sheet for each student which showed all assessments given by teachers, including facet assessments, symbols indicating preparedness to change and comments (see Appendix 4). Adjustments to the assessments were not made at this point.

Two major tasks were undertaken on the third school visit:

- the overall assessor worked through the process of synthesising the teachers' assessments into one overall assessment;
- the researchers distributed evaluation questionnaires and interviewed teachers about the process.

Overall assessment

Just prior to this third visit, each Overall Assessor was sent a document which outlined how they might best deal with the material that would be presented during the visit. Each Overall Assessor was expected to work through all of the assessments for that school noting, in particular, cases where the assessments given by teachers ranged over four or more points of the eight-point scale (eg. from 1 to 4). In some cases it was expected to be easy to contract discrepant marks to an acceptable range - for example in cases where outlying teachers had indicated (by means of symbols provided) that they would be prepared to have their marks shifted in a certain direction. Other cases would be more difficult, and it was expected that there might be some cases that would not be resolved.

As mentioned, Overall Assessors did not use a computer program for this exercise. The results were transferred to sheets resembling what had been planned to be displayed by the computer software.

The object of this activity was to see whether and to what extent the Overall Assessors could and would be prepared to use particular pieces of information in making an overall decision about the most appropriate result for an individual student.

Overall Assessors seemed to have less difficulty with this exercise than had been anticipated. There were some unresolvable cases, although sometimes this was because the exercise had to be completed on the day of the visit and there was not always time to hold discussions with teachers whose marks for particular students had been discrepant.

Questionnaire and interviews

A list was sent to schools which outlined the kinds of issues to be discussed with participants on the third visit. Most schools went to some trouble to ensure that participants could be available to take part in interviews at this exceptionally busy time of year. The outcomes of this data gathering are discussed in Part 3.

DEVELOPMENT OF REPORTING SOFTWARE

The project aims to develop a prototype computer program that can be used for four interrelated purposes:

- to enable teachers to enter their KC assessments for each student including levels and standardised descriptors;
- to enable the overall assessor to synthesise teachers' assessments into an overall level for each student;
- to produce the results on a report form; and
- to produce data about the assessments of use to the school.

Entry of data by teachers

It is proposed that teachers will undertake the same data entry as they undertook when completing assessments in the trial, but this will be done using the software package rather than pencil and paper. This will involve for each student:

- recording a point on an eight-point scale for each Key Competency;
- if needed, recording additional marks (eg. ?) to indicate the degree of certainty of their opinion;
- marking an asterisk to indicate that they wish to make a comment, then recording that comment on their perspective of the student or a comment that can be integrated into the report on the student that will be produced;
- nominating and marking a facet descriptor for each Key Competency that is the typical and distinctive strength of the student.

Synthesis by the overall assessor

The levels given and descriptors chosen by each teacher will be merged into one file and the program will be used to:

- calculate an algorithmic outcome on each Key Competency for each student, taking into account the levels recorded by different teachers (see Appendix 8);
- indicate whether the range of levels is irreconcilable (and should be reported only after consultation with particular teachers);
- review the descriptors identified by individual teachers and determine the most appropriate descriptor for each student.

The overall assessor will review the range of levels, the codes and comments given by teachers, and the algorithmic outcome calculated by the program to give each student a level for each Key Competency.

Reporting of levels

The program will be used to produce reports indicating levels for each Key Competency and typical descriptors on each Key Competency for each student. It will also be possible to produce a school descriptive statement in keeping with the model report (see Appendix 10).

Data for the school

The program will indicate if certain teachers are 'lenient' in their marking and if others are 'harsh'. These data may be useful when the overall assessor is producing a level for a student.

Descriptive statistics will be produced so that schools will be able to compare assessments across the Key Competencies, across teachers and across subject areas.

Part 3 Participants' Views of the Assessment Framework and Procedures

This Part of the report outlines the processes undertaken to gather reactions from participating teachers to the assessment framework developed by the project, the assessment processes trialled and the Key Competencies themselves.

The process undertaken

Teachers' perceptions of the trial exercise are a crucial part of the project. It was decided to seek their opinions by means of a questionnaire (see Appendix 5) and interviews or small group discussions. The face to face contact provided by an interview or discussion was favoured. This was thought to encourage points of view which might not have been elicited from the pre-conceived items of a questionnaire and to enable the follow through of leads from other comments. On the other hand, it was thought that teachers might feel inhibited by the presence of the researchers if they wished to make criticisms. Thus a questionnaire was used to gather some quantitative feedback and to provide a terrain which might be less threatening for teachers if they wished to make strong criticisms.

The data discussed here are from face to face discussions with 42 teachers and questionnaire responses from 58 teachers who had participated in the assessment process (a total of 110 teachers took part in the trial exercise). Most of the teachers who were interviewed also completed questionnaires. As similar issues were covered in the questionnaire and in interviews, the following discussion focuses on these issues, referring to data from both sources, as the questionnaire data and interview responses are considered to be mutually illuminating. During this evaluation procedure, the researchers spoke personally to about three-quarters of the participants. The main aim here is to indicate impressions gained from these encounters. The table below indicates the main subject areas taught by these teachers.

MAIN SUBJECT AREA	Percentage of teachers (questionnaire) N=58	Percentage of teachers (interview/ discussion) N=42
English/Humanities/Social Sciences	34.5	35.7
Mathematics/Science	31.0	26.2
Business/Info tech/Accounting	15.5	14.3
Performing arts/Visual arts/Phys ed/Home eco	15.5	16.7
'Other' (includes careers teacher/special ed teacher)	3.4	7.1

The Key Competencies in overview

This project aimed to explore strategies for assessing and reporting the Key Competencies rather than to seek opinions from teachers about the perceived value of such a conception. It seemed useful, however, to spend a little time discussing the Key Competencies as a conception in case teachers' views influenced their ability to take part in the assessment process. It is important to remember that some teachers participating in this project had barely heard of the Key Competencies prior to their involvement. As can be seen below, some thought them valuable and others were critical. Regardless of their stance in these matters, however, it was found that teachers were able to participate in the assessment strategies outlined.

Interviews provided an easier forum for discussion of the conception of a group of employment-related Key Competencies than did the questionnaire. The following discussion indicates the range of views of teachers participating in the trial. Teachers who were favourably disposed to the notion of Key Competencies tended to stress the role they can have in broadening the school curriculum or else to focus on the fact that they provide important goals which can be achieved by students who are usually weak academically.

Role of the Key Competencies in broadening the school curriculum

Typical comments on the broadening role were:

> *We've tried so many times to break the subject basis of the high school, and our training continually forbids it! What it may take is the superimposition of another system on top. Teachers can continue to do what they're doing, but they'll see that there are other things that need doing outside their subject area... I see the Key Competencies as a nice wedge into the future, into changing things - away from chalk and talk and cells and bells!*

and

> *It asked me to identify things in students that I don't usually consider when I assess them. I found it really valuable to think about other aspects.*

or looking more from the student's viewpoint:

> *It gives a broader understanding of what the kid's about. I suppose that's what we should be looking at rather than putting them in little boxes according to subject areas.*

Role of the Key Competencies for students with low academic achievement

Teachers who drew attention to the positive role Key Competencies can have for students who are usually low achievers in academic subjects said:

> *I think we do a lot of this already but we probably wouldn't report on it. It's a more positive way of reporting. Even kids who aren't good academically can achieve a medium or high in some of these [competencies]. One of our integration students was getting 4s and 5s. That was interesting.*

and

> *A report showing Key Competency 7 and 8 - even 6, would be especially useful for students with low academic ability but good people skills and good personalities.*

I've noticed this tendency with some students who are weak academically. Not all businesses want everybody Solving Problems!

Criticisms of the Key Competencies

Some teachers, however, were nervous about the 'human capital' emphasis that they saw in a conception such as the Key Competencies. Interestingly, teachers who made the following two comments still spoke favourably about the exercise in general and seemed happy to co-operate with it, but they had underlying concerns.

> *I resent this convergence between conventional and vocational education. I think it's a rort. Employers have monopolised what we're doing in schools, and this is one further step. It will be a gradual takeover which will continue until general education is decimated by employers. Employers have the ear of the government.*

> *It bothers me that education could become dominated by Key Competencies. And it bothers me because it doesn't address a purpose of education where people make meanings out of their lives and so forth.*

This last teacher was aware that the Key Competencies are 'employment related', but seemed concerned that they could be interpreted as having a broader role, which would be undesirable. This seems to be a fairly commonly held concern - that the Key Competencies are supposed to represent a total curriculum, or 'life', rather than generic skills relevant to the work place.

Visual and Performing Arts teachers felt that their areas should be represented in the Key Competencies. This sometimes seemed to be a general concern that employment related competencies should address issues such as aesthetic awareness or ability to be creative. In other cases it seemed to be a mistaken notion that the Key Competencies should embrace all 'important' curriculum areas.

> *There's not much creative - I suppose problem solving is a little bit. Not much room for creativeness.*

> *They do seem to overlook the performing arts areas.*

A few teachers complained about overwork, and an infrequently expressed view was that teachers are stressed by the amount of change that has taken place recently: they want to be left with the *status quo*, rather than having to adjust to more new approaches.

> *In a nutshell, teachers have had a gut full! They [the government] should probably spend more of their money on actually creating a few jobs, rather than trying to get new ways of making sure that employers know what our schools are about. These kids have skills, and I think they're reported reasonably well now.*

This teacher was, nevertheless, willing to take part in the trial and seemed to use the opportunity to voice a general concern rather than one directed specifically towards this project.

Clarity of the Key Competencies

As mentioned in Part 2, the Key Competencies were mapped by means of facet descriptors to try to ensure as far as possible that teachers in the trial were assessing the same constructs.

The following discussion of issues of clarity address both the overall conception of Key Competencies and test the clarity with which they were mapped for this project.

Questionnaire data indicate that 13.8 per cent of the teachers found the Key Competencies 'very clear', 53.4 per cent 'quite clear' and 24.1 per cent 'clear'. Only 44 per cent of teachers of Business Studies/Computing thought the Key Competencies as a conception to be very clear or quite clear compared to 74 per cent of the Maths/Science teachers and 80 per cent of the English/Humanities teachers. The four teachers classified as 'other' (which included careers teachers and a school principal) seemed to have a more negative view of the conception. Numbers representing subject groups are low, however, so it is unwise to place much weight on these observations

The Key Competency found to be most unclear or unintelligible was Using Cultural Understandings, which was mentioned by 32.8 per cent of the questionnaire respondents. This is not surprising given that this has been the most problematic of the Key Competencies and subject to prolonged debate. Some teachers who were interviewed felt that this Key Competency could be clarified if it had a stronger 'social issues' emphasis, for example:

> *May be [it should be more to do with] social awareness. You're really after how in tune students are with what's going on around them. It's more about social awareness than cultural understanding.*

Others found it difficult to define and therefore difficult to assess:

> *It's in the kid's head. You don't necessarily see it.*

Other Key Competencies considered unclear by more than one respondent to the questionnaire were Communicating Ideas and Information - Oral Communication (N=3) and Using Mathematics (N=4). Just over 15 per cent responded that 'none' were unclear and 27.6 per cent omitted the question which could have been an indication that they felt the Key Competencies were clear and intelligible.

Least useful Key Competencies

Key Competencies that two or more questionnaire respondents thought should be deleted from the group were:

- Using Mathematics (8.6 per cent) - even a Maths teacher who was interviewed said: *'I don't think you need this Key Competency'* - the implication being that these skills are covered elsewhere.
- Using Cultural Understandings (12.1 per cent)

and

- Using Technology (3.5 per cent)

The table below outlines reasons given (by questionnaire respondents) for deleting particular Key Competencies from the group.

KEY COMPETENCY TO DELETE	REASONS GIVEN FOR RECOMMENDING DELETION
Using Mathematical Ideas and Techniques	already covered with problem solving not relevant to all learning areas unclear (as a conception) not all teachers can judge this area

Using Cultural Understandings	not very relevant to employers too open-ended too difficult to assess
Using Technology	too difficult to assess 'does not relate to my area'

There may be a relationship between teachers' main subject areas and desire to delete particular Key Competencies from the group, although it must be remembered that numbers of teachers representing each subject in this sample are small. The table below indicates the subject areas of teachers who wanted to delete Key Competencies.

KEY COMPETENCY	MAIN SUBJECT AREA OF TEACHERS WHO WANTED TO DELETE IT	MAIN SUBJECT AREA OF TEACHERS WHO DID NOT INDICATE THAT THEY WANTED TO DELETE IT
Using Mathematical Ideas and Techniques	English/Humanities (15%) Business/Info Tech (22%)	Maths/Science Performing Arts/ Visual Arts
Using Cultural Understandings	English/Humanities(5%) Maths/Science (33%)	Performing Arts/ Visual Arts Business/Info Tech
Using Technology	Maths/Science(13%) Performing Arts(25%)	English/Humanities Business/Info Tech

In all ten schools involved in the trial exercise it was found that English/Humanities teachers (and indeed, in most cases teachers of any subject area that was not Maths) were anxious about assessing Using Mathematical Ideas and Techniques. This was apparent even in subject areas such as Science and Accounting where it might have been expected that teachers would feel comfortable in the Mathematics area. It is clear both from these evaluative data and from analysis of teachers' assessments that the Key Competencies whose conceptions are still somewhat problematic are the three referred to in earlier tables: Using Mathematical Ideas and Techniques, Using Cultural Understandings and Using Technology.

Suggested additions to the Key Competencies

In the questionnaire, teachers were invited to suggest areas that could be added to the group of Key Competencies to make them more satisfactory as a conception of employment-related competencies. The following were suggested by two or more teachers:

- Punctuality/time management (6.9 per cent)
- Logical processes (3.5 per cent)
- Creative/imaginative processes (3.5 per cent)
- 'relate more to subject areas' (3.5 per cent)
- broaden Using Cultural Understandings (3.5 per cent)

In relation to this question 13.8 per cent of the teachers wrote 'nothing' and 41.4 per cent omitted the question, which could indicate a belief that nothing needed to be added.

Summary of teachers' views of the Key Competencies

Questionnaire data indicate that overall teachers found the Key Competencies presented in the project satisfactory as a conception of employment-related competencies. Just over 17 per cent found them 'very satisfactory', 48.3 per cent 'quite satisfactory' and 24.1 per cent 'satisfactory' and teachers with both positive and negative views generally found the actual concept of Key Competencies clear and intelligible.

Facets of the Key Competencies

Facet descriptors were written to help define the Key Competency and thus facilitate its assessment. It was important to find out whether these facets:

- identify aspects of the activity that can be usefully and meaningfully separated from other aspects of the activity
- identify activities that are sufficiently like each other to remain a single and coherent competency
- give teachers a clear basis for making an on-balance overall judgement

From the questionnaire data gathered, there will follow a summary which indicates for each Key Competency:

- facets thought to be unclear by five or more respondents to the questionnaire
- facets that two or more questionnaire respondents thought should be deleted
- new facets that two or more questionnaire respondents felt should be added
- an indication of whether the KC facets were thought to be not sufficiently different or too different.

Where appropriate, questionnaire data will be illuminated by comments from interviews.

Collecting, Analysing and Organising information

Facets thought to be unclear by five or more respondents	nil
Facets that two or more respondents thought should be deleted	Facet B: analysing and evaluating ideas and information (N=4)
New facets that two or more respondents felt should be added	separate analysing and evaluating (Facet B) (N=4)
Are the facets sufficiently but not too different from each other?	Too different: N=1 Different and coherent: N=42 Not sufficiently different: N=9 Omit: N=6

Backing up the comments about Facet B, teachers who were interviewed pointed out that the terms 'analysing', 'synthesising' and 'evaluating' may be interpreted in different ways and depending on the interpretation they may be considered to be more or less demanding. The facet descriptors were not intended to be hierarchical, but a number of teachers thought that this was the intention with the mapping of this Key Competency. Some even said that it reminded them of Bloom's taxonomy.

Communicating Ideas and Information - Oral Communication

Facets thought to be unclear by five or more respondents	Facet A: speaking in less formal situations (N=6) Facet C: using speech to interact with other people (N=7)
Facets that two or more respondents thought should be deleted	Facet C: using speech to interact with other people (N=10)
New facets that two or more respondents felt should be added	ability to communicate with adults (N=2)
Are the facets sufficiently but not too different from each other?	Too different: Nil Different and coherent: N=38 Not sufficiently different: N=11 Omit: N=9

Many of the interview comments suggest that Facets A and C are very similar. For example: *'I wasn't sure what was meant by less formal situations. Facets A and C seemed similar.'*

Communicating Ideas and Information - Written Communication

Facets thought to be unclear by five or more respondents	Facet B: writing confidently and fluently (N=5) Facet C: using writing to explore ideas and information (N=7)
Facets that two or more respondents thought should be deleted	Facet B: writing confidently and fluently (N=3) Facet C: using writing to explore ideas and information (N=5)
New facets that two or more respondents felt should be added	using a variety of text types (N=3) writing for specified audiences (N=2)
Are the facets sufficiently but not too different from each other?	Too different: Nil Different and coherent: N=36 Not sufficiently different: N=16 Omit: N=6

In interviews some English teachers commented that they saw a big gap in the demands made by some of the facets. For example, one teacher commented: *'Using writing to explore ideas and information is a very high level thing to do. Not many students do that.'* This is not a serious concern - it possibly indicates that not many teachers will pick this as a student's strongest facet.

Using Mathematical Ideas and Techniques

Facets thought to be unclear by five or more respondents	Facet A: Using Mathematical Ideas and Techniques to analyse issues (N=8) Facet B: understanding mathematical principles and procedures (N=8) Facet C: applying mathematical knowledge to problems (N=7)
Facets that two or more respondents thought should be deleted	Facet B: understanding mathematical principles and procedures (N=2)
New facets that two or more respondents felt should be added	able to use basic mathematical tools (N=2)
Are the facets sufficiently but not too different from each other?	Too different: nil Different and coherent: N=35 Not sufficiently different: N=6 Omit: N=17

Using Cultural Understandings

Facets thought to be unclear by five or more respondents	Facet A: comprehending and interpreting social issues (N=9) Facet B: developing a coherent position on social issues (N=12) Facet C: Using Cultural Understandings to achieve goals (N=28)
Facets that two or more respondents thought should be deleted	Facet C: Using Cultural Understandings to achieve goals (N=12)
New facets that two or more respondents felt should be added	suggestion that examples should be given for clarification (N=2) - this suggestion was also made by other respondents applied in general to the KCs include attitudes to different cultures (N=3)
Are the facets sufficiently but not too different from each other?	Too different: N=1 Different and coherent: N=33 Not sufficiently different: N=11 Other: N=1 Omit: N=12

Several teachers who were interviewed, backed up the feeling of lack of clarity in Facet B, for example:

> *There are not enough situations I see kids in to know whether they have developed a coherent position... they are all alike... the only cultural understanding they have is of American TV culture. They don't reflect on the bigger picture and why is this happening - I don't think kids do that. Perhaps it could be defined as: Do they understand the influence of cultures on their behaviours?*

and

> *A coherent position can be very positive or very negative!*

This last comment suggests that the attempts to make Using Cultural Understandings cognitive rather than attitudinal may not have been entirely successful, as the comment implies a belief that the teacher is to assess whether a student has a 'good' or a 'bad' position.

Some of the difficulty in defining Facet C was expressed in this comment:

> *Everybody has got different goals. How can you assess whether they are Using* ***Cultural*** *Understanding to achieve their particular goals?*

another teacher explained how he interpreted this facet:

> *I thought it was their [the students'] interpretation of what is happening around them. If they felt strongly about something would they do anything about it?*

Solving Problems

Facets thought to be unclear by five or more respondents	nil
Facets that two or more respondents thought should be deleted	nil
New facets that two or more respondents felt should be added	shows a reasoned and methodical approach (N=2)
Are the facets sufficiently but not too different from each other?	Too different: nil Different and coherent: N=51 Not sufficiently different: N=4 Omit: N=3

Although questionnaire respondents did not reveal concerns with the facets, a few issues were mentioned by teachers who were interviewed. One Maths teacher had a lot of difficulty with the notion of rewarding independence and initiative or focus and persistence if a student who was high on these skills ended up with an incorrect answer:

> *Solving Problems was the one I had most difficulty with. In Maths, that's more or less what they do...but 'focus and persistence' is getting an answer in the end, whether they're good at it or not. They might be independent, but not come up with an answer. They should be coming up with something reasonable at the end of it. They might show all these facets, but not come up with something reasonable in the end.*

The issue of transferability - whether a skill displayed in one context is necessarily transferred to another context - is a significant issue in the consideration of assessing generic or cross-curricular skills. It was raised by this history teacher when commenting on Solving Problems:

> *If you're looking at evidence and working from historical documents, that's what I based what I did on. But it's a very small sample of a child's ability to solve problems. I wouldn't like to generalise from that about their general problem solving ability.*

To some extent this concern is addressed by the fact that a student's final report could be compiled from the judgements of a number of teachers. But this issue also relates to how teachers approached global impression judgements, which is discussed later in this section.

Using Technology

Facets thought to be unclear by five or more respondents	Facet C: can understand technological systems (N=7)
Facets that two or more respondents thought should be deleted	Facet B: is thorough and persistent in dealing with technology (N=2) Facet C: can understand technological systems (N=5)
New facets that two or more respondents felt should be added	request that the areas of technology be specified (N=2)
Are the facets sufficiently but not too different from each other?	Too different: nil Different and coherent: N=41 Not sufficiently different: N=7 Omit: N=10

The facets for this Key Competency were written with the intention of giving an attitudinal focus. This was an attempt to overcome issues of equity, such as opportunities to use technology and to try to keep the focus cross-curricular, rather than looking at technologies in relation to particular subject areas. But this did not always seem to be picked up by the teachers in the study. Some teachers who were interviewed indicated that they assumed that 'technology' meant computers and one teacher suggested that this is how it will be interpreted in spite of efforts to give a more general definition:

> *The more you try to define some of these things, the harder it is to grasp. What is technology? Is it computers? The employers aren't going to bother to read through this stuff to see what the definition is. They'll probably assume it's computers. 'Doesn't seek computer experiences - oh, well - Don't want him!'*

Another teacher seemed to assume that this Key Competency referred to sophisticated kinds of technology when she said:

> *How can you say that our kids **seek** technological experiences? There's nothing for them to seek! It starts with an assumption [that the school has 'technological' equipment].*

Planning and Organising Activities

Facets thought to be unclear by five or more respondents	nil
Facets that two or more respondents thought should be deleted	Facet C: seeks planning and organisational experiences and challenges (N=6)
New facets that two or more respondents felt should be added	nil
Are the facets sufficiently but not too different from each other?	Too different: nil Different and coherent: N=49 Not sufficiently different: N=4 Omit: N=5

Working with Others and in Teams

Facets thought to be unclear by five or more respondents	nil
Facets that two or more respondents thought should be deleted	Facet A: adapts to group expectations (N=6)
New facets that two or more respondents felt should be added	nil
Are the facets sufficiently but not too different from each other?	Too different: N=1 Different and coherent: N=52 Not sufficiently different: N=4 Omit: N=1

Some of the difficulty with Facet A was explained by the following teacher when she was interviewed:

> *I found I was applying 'adapts to group expectations' to kids who really don't have much self esteem and just want to go along with the mob. We have numbers of kids who just don't work in groups.*

Another teacher who was interviewed said that he didn't understand what was meant by 'adapts to group expectations'.

There is a clear indication that the mapping of the Key Competencies through the facets has been successful. It was heartening when one teacher said in interview that he would have been 'lost' without the facets. Some 'fine tuning' to facet descriptors may be needed. The table below outlines areas where modifications may be needed.

Facet descriptors where modification may be needed

KEY COMPETENCY	COMMENTS ON FACET DESCRIPTORS
Collecting Analysing and Organising Information	different interpretations of analysing and evaluating (Facet B)
Communicating Ideas and Information - Oral	confusion about meanings of 'less formal' and 'more formal' - similarity between Facets A and C
Communicating Ideas and Information - Written	Facets B and C thought to be unclear - suggestion that Facet C is difficult
Using Mathematical Ideas and Techniques	all three Facets thought to be unclear, but this probably relates more to teachers' overall difficulty with this Key Competency
Using Cultural Understandings	as for Using Mathematical Ideas
Solving Problems	all Facets clear
Using Technology	difficulties with Facets B and C - but again, teachers have difficulty with the overall Key Competency - they tended not to recognise the purpose of the 'attitudinal' interpretation
Planning and Organising Activities	some teachers are unhappy with Facet C
Working with Others and in Teams	some difficulty with interpretation of Facet A

Teachers' views of the Facets in light of the main subjects taught

Are there notably different perceptions of Key Competencies according to the main subject areas that are taught? The table below indicates that there are few striking differences when considering the numbers of teachers who found the Facets sufficiently different, but not too different from each other. The table below gives the numbers (with percentages in brackets) of teachers in specified subject areas who indicated that the Facets of a Key Competency were 'different and coherent'.

Indication, according to subject taught, of whether the Facets were sufficiently different and coherent

MAIN SUB-JECT AREA	KC1	KC2A	KC2B	KC3	KC4	KC5	KC6	KC7	KC8
Eng/ Hum N=20	14 (70.0)	15 (75.0)	16 (80.0)	14 (70.0)	12 (60.0)	17 (85.0)	12 (60.0)	19 (95.0)	20 (100.0)
Maths/ Science N=18	15 (83.3)	11 (61.1)	9 (50.0)	13 (72.2)	9 (50.0)	16 (88.9)	12 (66.7)	16 (88.9)	16 (88.9)
Bus N=10	5 (50.0)	7 (70.0)	4 (40.0)	4 (40.0)	5 (50.0)	9 (90.0)	7 (70.0)	8 (80.0)	9 (90.0)
Arts N=7	6 (85.7)	4 (57.1)	5 (71.4)	3 (42.9)	4 (57.1)	5 (71.4)	4 (57.1)	4 (57.1)	5 (71.4)
Other N=4	3 (75.0)	2 (50.0)	2 (50.0)	2 (50.0)	1 (25.0)	3 (75.0)	3 (75.0)	3 (75.0)	3 (75.0)

(N= the number of teachers of that subject area making responses to the question: Are the Facets sufficiently but not too different from each other?)

Given the small numbers representing each subject area, the main areas worth comparing are Maths/Science and English/Humanities. On the whole there is a good level of agreement. It seems that English/Humanities teachers were a little less happy with the Facets for Collecting, Analysing and Organising than were Maths/Science teachers but that English/Humanities teachers were more satisfied with the Facets for both parts of Communicating Ideas and Information. Given that Using Mathematical Ideas and Techniques was seen as a problematic Key Competency - especially by English/Humanities teachers, it is interesting to note that similar percentages of English/Humanities and Maths/Science teachers (around 70 per cent) found the Facets to be sufficiently different and coherent. The other somewhat problematic Key Competencies - Using Cultural Understandings and Using Technology - do have slightly lower percentages with the Facets.

Summary of participants' views of the Facets

Overall it seems that the notion of Facets has worked and that the Facets developed were generally sufficiently but not too different from each other, so that they were perceived as defining a coherent concept. Teachers' responses indicate a remarkable degree of agreement and acceptance of the mapping of Key Competencies presented. For example, 90 per cent of the teachers who responded to the questionnaire indicated that the Facets for Working in Teams were 'different and coherent' - in other words the Facets gave a complete definition of the Key Competency but each expressed a different aspect of it. There was also a very high level of agreement about Planning and Organising (84 per cent) and Solving Problems (88 per cent). Some Key Competencies appeared to be a little more problematic, but closer inspection suggests that this was often due to the need for some 'fine tuning' to the wording of Facets. This is the case with Communicating Ideas and Information - oral (66 per cent) and written (62 per cent) and Collecting, Analysing and Organising (72 per cent). The conceptions of Using Mathematics, Using Cultural Understanding and Using Technology may require further thinking although in all cases more than 50 per cent of the respondents indicated that the Facets were 'different and coherent'. This shows a remarkable degree of consensus and indicates that the definitions given are largely workable.

It is particularly interesting to note that the less cognitive or more attitudinal Key Competencies such as Planning and Organising and Working in Teams were so well received. In the past educators have tended to shy away from assessing attitudes, yet it seems that these Key Competencies operated even better than the Key Competencies more aligned with traditional curriculum areas - such as Using Mathematics and Written and Oral Communication.

THE FRAMEWORK AND THE PROCESS

Use of the framework

Fifty one of the fifty eight teachers who completed questionnaires found that the level descriptors were 'very useful' or 'quite useful'. In this context, 17.2 per cent of the teachers commented that they would find exemplars helpful. A few (10.3 per cent) wanted more options, but on the other hand some (3.5 per cent) thought that there were too many and found the extra symbols - arrows etc. confusing. A small proportion (3.5 per cent) suggested that the subjects studied by a student should be taken into account when assessing the Key Competencies.

In interviews teachers did not make much reference to the assessment framework. There was only one suggestion about changing the format - one teacher asked that the 1 - 8 scale be reversed with 8 (the highest possible score) at the top. General comments suggested that the assessment procedure was easier than expected:

> *It wasn't as hard as I expected.*
>
> *It's really just a different way of giving information that you already have.*
>
> *I did twenty or more kids and the more I did, the easier it became to use.*

Unfavourable general comments were largely concerned with the imposition on already stretched schedules:

> *I have a problem with it overall because it's* ***another*** *thing that teachers have to do!...If I were teaching Accounting and had to comment on all those things, I'd come back to my initial problem [too much work]....It would be OK if you didn't have to assess every Key Competency.*
>
> *It did take a long time. I had to keep going back and checking.*

Time taken to complete assessments

In interviews most teachers said that they took about 10 minutes to assess each student on all the Key Competencies once they had become familiar with the procedure. One teacher took 20 minutes per student. Some teachers saw 10 minutes per student as an excessive amount of time - they may have been considering this in relation to the project being an imposition on their usual work. Others felt that 10 minutes was reasonable and cost effective.

Views about making global impression judgements

The assessment procedure required teachers to make global, on-balance judgements of students' Key Competencies. Teachers were asked whether they found this an easy task. More teachers found it easy than difficult:

- very easy: 12.1 per cent
- quite easy: 48.3 per cent
- quite difficult: 22.4 per cent
- very difficult: 3.5 per cent
- other (eg. a combination): 6.9 per cent
- omit: 6.9 per cent

It was thought that there might be a relationship between subject area and perceptions about the difficulty of making global judgements - for example that English and Humanities or Arts teachers might find the process easier because they are more inclined towards making these kinds of judgements. The table below shows teachers' responses to the question: 'How did you find making global judgements about the Key Competencies performance of the students you assessed in this trail?' according to their main subject area.

Teachers' perceptions of the difficulty of making global impression judgements

MAIN SUBJECT TAUGHT	VERY EASY	QUITE EASY	QUITE DIFFICULT	VERY DIFFICULT
English/Humanities N=18	3 (16.7)	11 (61.1)	4 (22.2)	
Maths/Science N=16	2 (12.5)	8 (50.0)	5 (31.3)	1 (6.3)
Business N=7		2 (28.6)	4 (57.1)	1 (14.3)
Arts N=5	1 (20.0)	4 (80.0)		
Other N=4		3 (75.0)	1 (25.0)	

(N= the number of teachers of that subject area responding to the question)

There is a very slight tendency for Maths/Science teachers to find the process more difficult than English/Humanities and an indication that perhaps Visual and Performing Arts teachers find making these kinds of assessments easier than do teachers of other subjects, but numbers are too small to enable firm conclusions to be drawn.

Comments from the questionnaire about the process of making global judgements being easy were:

- most teachers are used to assessing this way (N=3)
- it's easy providing you know the students well (N=8)
- it was easy to internalise and apply the standards (N=5)
- it became easier with practice (N=2)

Comments from the questionnaire about the process of making global judgements being difficult were:

- there was not enough time to be accurate (N=2) -(from discussions this probably meant that teachers felt they didn't have sufficient time to observe the students)
- the course isn't structured to fit this (N=4)

- it's difficult if you are not familiar with all the students (N=6)
- uncomfortable with the subjectiveness (N=5)
- difficult with a subject specific KC (eg. Using Maths) when the teacher isn't familiar with the area (N=2)
- the framework was difficult to use (N=3)

Comments in interviews suggested that teachers who were comfortable about making global impression judgements saw themselves as experienced professionals whose judgements should be respected. As one teacher said:

> *People [teachers] should be proud of their professional judgement. Employers respect it.*

These teachers indicated that they are experienced at making these sorts of judgements:

> *I can tell you within the first six weeks of school what a kid's going to get at the end of the year. You know. You might be wrong with about three kids [out of the class]. Of course, it's important not to be wrong with those three kids. But I feel comfortable with it.*

This view was repeated by a teacher from another school (and state):

> *You're looking at a lot of professional opinion here. And in many cases it's spot on. Even with the HSC, without looking at the kids' results I can rank them. I'm usually only about one or two students out.*

Teachers sometimes summed up this professional judgement as a 'gut feeling', but they felt that they could trust this feeling:

> *You go ahead and do it and see what comes out of it. The goal is to be subjective.*

> *You've got to trust - you've got to take a deep breath and say, well, bugger it, that's what I think! You've got to trust your gut feelings a fair bit.*

Teachers who felt uncomfortable about making global judgements seemed to feel insecure because of a lack of back-up support such as test results. For example:

> *I felt I didn't have enough back-up data to make judgements outside what I normally do.*

> *I felt very uncomfortable about it. I suppose I could have gone through mark books for evidence and so forth - but I didn't have time. It's a calculated guess.*

It seemed that these teachers were very 'subject bound' - in interviews they talked more about teaching their subject area than about the students in general terms.

> *There's a lot of value judgement when it's not mark related. I might get on better with some students than with others... If a student doesn't like a particular subject, this will be reflected.... There was nothing to back you up. Some kids you only see for one subject. And some of this takes it outside the classroom. Some you don't know except in the classroom.*

It is clearly essential that teachers know their students well if they are to make confident global impression judgements.

In interviews teachers with negative feelings about making global impression judgements often became less anxious when they realised that in a real situation they would have at least a year to think about students' performance in the Key Competency areas and it was helpful to be reminded that they are not the only person judging a student in a particular area.

In relation to this it was interesting to observe how overall assessors went about the task of adjusting those assessments where teacher agreement was outside an acceptable band (such as 1 -3). Some assessors relied quite heavily on their own knowledge of students and teachers whereas others went through a form of averaging process - even although this had not been recommended. The first method seemed to rely quite heavily on a form of professional 'gut feeling'.

Teachers were asked whether they found the assessment cost effective. Unfortunately some (12.1 per cent) were unsure as to what was meant by 'cost effective'. The most popular response was that the exercise was 'quite cost effective' (48.3 per cent). Eight teachers (13.8 per cent) thought it 'not very cost effective', three (5.2 per cent) thought it 'not cost effective', but two (3.5 per cent) thought it 'very cost effective'.

Finally teachers were asked whether they felt that an overall assessment made from a number of teachers' judgements would give useful information to employers. Teachers clearly felt that the material would be useful:

- very useful: 32.8 per cent
- quite useful: 46.6 per cent
- not very useful: 6.9 per cent
- of little use: 6.9 per cent
- other, or omit: 6.9 per cent

Conclusion

The evaluative data discussed do indicate that teachers are able to use the assessment procedure and that most found it easier than they had expected. Teachers who were anxious about the reliance on global impression judgements tended to be happier when they were reminded that the trial exercise was only experimental and that in a real situation they would have much more time to become familiar with students' performances in the Key Competency areas.

Teachers found that the overall assessment framework was clear and they were able to use it with remarkably little instruction. One teacher found that it took only ten minutes to fully explain the procedure to a colleague who had been unable to attend the inservice session.

As might be expected, some teachers believe that the Key Competencies may provide an opportunity for less academically able students to achieve. A clear majority of teachers who completed the questionnaire believe that the Key Competencies will be useful to employers.

2 Part 4 The Key Competencies Levels Trial Data and Analysis

The following discussion outlines the quantitative data obtained from the assessment trial procedure undertaken in this project. The data and the analysis were designed to address two basic issues. The first issue is the characteristics of the outcomes produced by the assessment procedures, and the second is the reliability, in terms of the inter-rater agreement or consistency, of the assessments made by the participating teachers.

In other words, the analysis deals with the following issues.

- How did the participating teachers use the Key Competencies score scale?
- What sorts of results did they produce?
- Did the participants discriminate among the students?

- How much did the participants agree with each other in the way they used the score scales and the way they assessed individual students?

- What do the data say about the Key Competencies as a set of constructs?

Distribution of the assessments

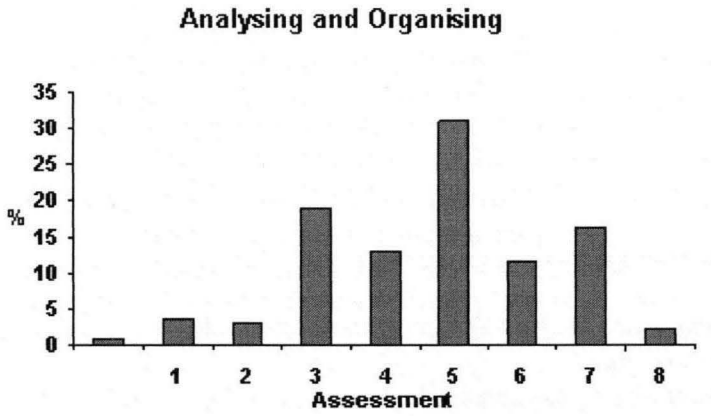
Figure 1 gives histograms of the percentage for each score point given by all participants for all Key Competencies. In general these histograms show that teachers discriminated quite significantly among the students. Although comparatively small numbers of students were given scores of 1, 2 and 8, most students were spread between 3 and 7, an effective range of five points.

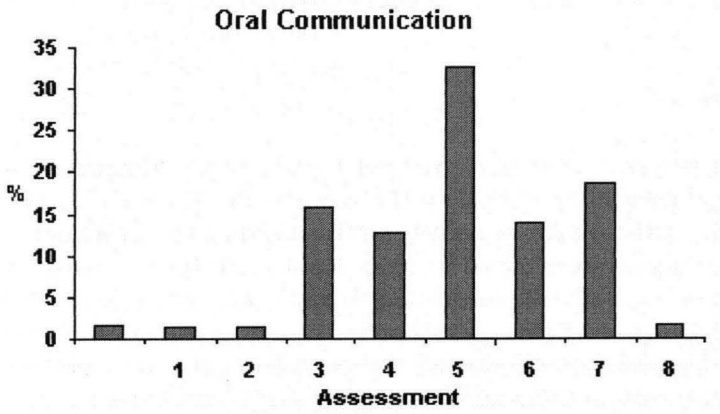
As was expected, the division of three levels into two sub-levels produced the 'saw-toothed' distribution to be seen in the histograms. In general scores of 3, 5 and 7 were more commonly given than 4 and 6, and in all cases 5 was the modal mark.
Apart from the dips at 4 and 6, the distribution shown in the histograms is approximately normal. The dips result from the presentation of the scale as having 3 levels and each divided into 2 sub-levels. This has the impact of encouraging assessors to place students just into a level, that is place them at the bottom sub-level. If the assessors are tempted to give the student the top of the level, they are also tempted to move the student into the next level, and hence the dips at scores 4 and 6.

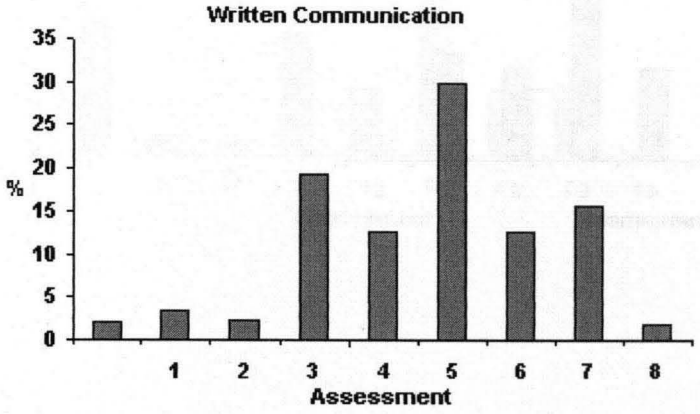
The histograms show that teachers did not inappropriately bunch the students, and they were making meaningful discriminations among them.

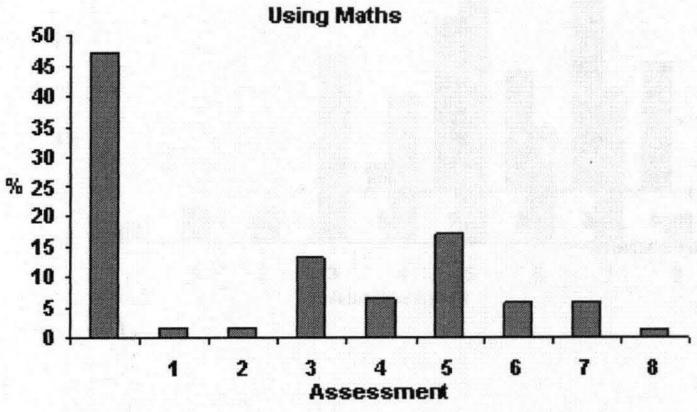
With the exceptions of Maths, Technology and Cultural Understandings, teachers were able to produce assessments of most of the Key Competencies. Assessments of Maths were not produced by teachers in 47% of cases, in 32% of cases assessments of Using Technology were not produced, and Cultural Understandings was not assessed for 15% of cases. The number of assessments and the distributions of the scores were quite consistent for the other Key Competencies.

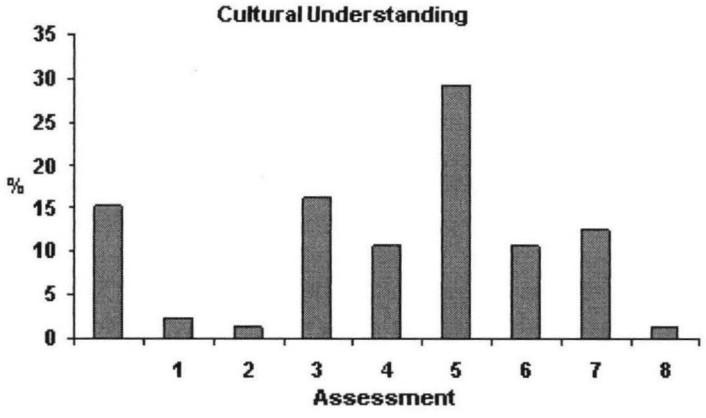
Figure 1 **Histograms of the percentage for each score point (or no score) give by all participants for all Key Competencies**

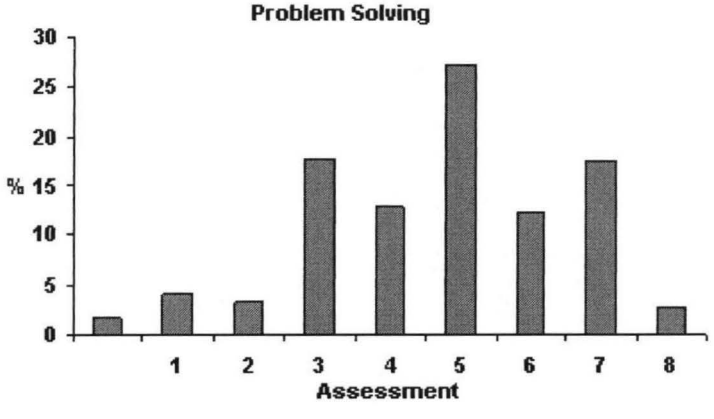


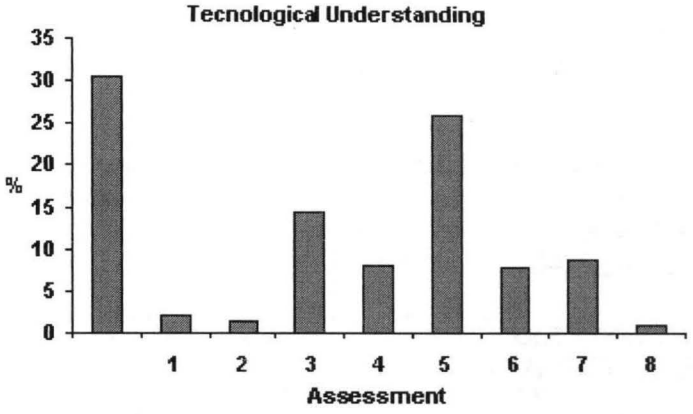



Written Communication
35
30
25
20
%
15
10
5
0
1
2
3
4
5
6
7
8
Assessment


Using Maths
50
45
40
35
30
% 25
20
15
10
5
0
1
2
3
4
5
6
7
8
Assessment


Cultural Understanding
35
30
25
20
%
15
10
5
0
1
2
3
4
5
6
7
8
Assessment


Problem Solving
30
25
20
% 15
10
5
0
1
2
3
4
5
6
7
8
Assessment


Tecnological Understanding
35
30
25
20
%
15
10
5
0
1
2
3
4
5
6
7
8
Assessment

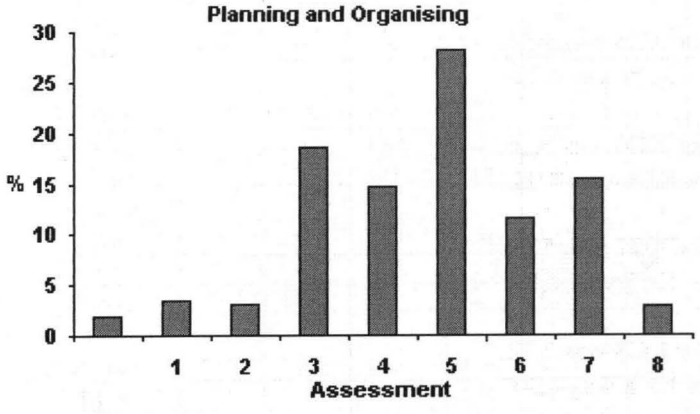

Planning and Organising
30
25
20
% 15
10
5
0
1
2
3
4
5
6
7
8
Assessment

Table 7 shows the summary statistics (average and standard deviation) for each school (numbered 1-10) and for all schools. The aggregated totals for each Key Competency in Table 7 show some differences between them. Oral C. has the highest average at 5.02, followed closely by Teamwork at 5.01. Maths is lowest at 4.57 followed by Technology at 4.69.

The school level data in Table 7 shows substantial differences in averages between schools.

- The highest school mean for KC1 was 5.25.
- The lowest school mean for KC1 was 3.94.

- The highest school mean for KC2a was 5.85.
- The lowest school mean for KC2a was 4.45.

- The highest school mean for KC2b was 5.38.
- The lowest school mean for KC2b was 4.25.

- The highest school mean for KC3 was 5.18.
- The lowest school mean for KC3 was 3.59

- The highest school mean for KC4 was 5.41.
- The lowest school mean for KC4 was 4.21.

- The highest school mean for KC5 was 5.54.
- The lowest school mean for KC5 was 4.23.

- The highest school mean for KC6 was 5.56.
- The lowest school mean for KC6 was 4.03.

- The highest school mean for KC7 was 5.63.
- The lowest school mean for KC7 was 4.10.

- The highest school mean for KC8 was 5.72.
- The lowest school mean for KC8 was 4.30.

Table 7 Summary Statistics for each school (1-10) and all schools

hool	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Aver	4.21	4.89	4.25	3.77	4.21	4.24	4.41	4.11	4.61
StDev	1.50	1.38	1.49	1.29	1.49	1.63	1.26	1.43	1.56
Count	173	175	174	83	173	174	115	175	173
2 Aver	4.73	4.96	4.79	4.52	5.00	4.66	4.64	4.72	4.82
StDev	1.37	1.21	1.33	1.30	1.23	1.54	1.23	1.45	1.41
Count	282	280	281	111	233	279	157	280	280
3 Aver	5.22	5.51	5.38	4.92	5.41	5.18	4.78	5.08	5.64
StDev	1.57	1.34	1.62	1.44	1.45	1.70	1.53	1.66	1.48
Count	117	117	117	71	113	116	87	113	113
4 Aver	3.94	4.45	4.27	3.59	4.22	4.23	4.03	4.10	4.56
StDev	1.88	1.76	1.73	1.67	1.80	1.94	1.63	1.94	1.78
Count	161	159	156	76	158	159	128	157	156
5 Aver	4.99	4.91	4.96	5.18	5.04	4.99	4.74	4.89	5.01
StDev	1.78	1.50	1.74	1.48	1.50	1.83	1.62	1.66	1.65
Count	117	115	115	73	103	101	61	115	115
6 Aver	5.14	5.06	5.06	4.55	4.85	5.04	4.92	5.05	5.18
StDev	1.56	1.62	1.65	1.51	1.76	1.63	1.43	1.59	1.59
Count	241	236	234	118	186	240	146	236	219
7 Aver	4.72	4.90	4.47	4.40	4.57	4.80	4.28	4.63	5.05
StDev	1.64	1.47	1.76	1.85	1.50	1.84	1.78	1.69	1.83
Count	156	155	155	92	121	158	116	153	155
8 Aver	4.65	5.07	4.66	4.30	4.63	4.76	4.35	4.61	5.01
StDev	1.76	1.50	1.76	1.63	1.52	1.76	1.70	1.77	1.75
Count	153	155	155	83	136	156	141	155	155
9 Aver	4.44	4.53	4.38	4.71	4.57	4.23	4.55	4.18	4.30
StDev	1.58	1.43	1.48	1.47	1.18	1.58	1.28	1.50	1.60
Count	132	130	126	73	69	132	91	132	133
10 Aver	5.25	5.58	5.18	5.18	5.33	5.54	5.56	5.63	5.72
StDev	1.37	1.29	1.36	1.40	1.17	1.34	1.00	1.31	1.20
Count	258	256	257	174	237	258	211	256	254
All Aver	4.76	5.02	4.77	4.57	4.81	4.80	4.69	4.76	5.01
StDev	1.62	1.48	1.61	1.57	1.52	1.71	1.50	1.65	1.61
Count	1790	1778	1770	954	1529	1773	1253	1772	1753

Table 8 shows the averages for each school ordered from highest to lowest for each Key Competency. School 10 had a higher average than all other schools in 6 of 9 Key Competencies. School 4 had a lower average than all other schools in 5 of 9 Key Competencies.

Table 8 also shows some differences between different Key Competencies averages within schools. School 5 was at or above the mean for all Key Competencies except Oral Communication. This school was also highest in Maths. School 7 was above the mean on Oral C, and well below on Maths and Technology. School 9 was below the mean for all Key Competencies except Maths. School 3 was within the highest 3 for all Key Competencies. These comparisons give schools information about the way their teachers see their students in comparison with other schools.

Table 8 The rank ordered means (highest to lowest) for all schools on each Key Competency

	KC1	Sch.	KC2a	Sch	KC2b	Sch	KC3	Sch	KC4	Sch
Highest	5.25	10	5.58	10	5.38	3	5.18	5	5.41	3
2	5.22	3	5.51	3	5.18	10	5.18	10	5.33	10
3	5.14	6	5.07	8	5.06	6	4.92	3	5.04	5
4	4.99	5	5.06	6	4.96	5	4.71	9	5	2
5	4.73	2	4.96	2	4.79	2	4.55	6	4.85	6
6	4.72	7	4.91	5	4.66	8	4.52	2	4.63	8
7	4.65	8	4.9	7	4.47	7	4.4	7	4.57	7
8	4.44	9	4.89	1	4.38	9	4.3	8	4.57	9
9	4.21	1	4.53	9	4.27	4	3.77	1	4.22	4
Lowest	3.94	4	4.45	4	4.25	1	3.59	4	4.21	1
Aver.	4.76		5.02		4.77		4.57		4.81	

	KC5	Sch	KC6	Sch	KC7	Sch	KC8	Sch
Highest	5.54	10	5.56	10	5.63	10	5.72	10
2	5.18	3	4.92	6	5.08	3	5.64	3
3	5.04	6	4.78	3	5.05	6	5.18	6
4	4.99	5	4.74	5	4.89	5	5.05	7
5	4.8	7	4.64	2	4.72	2	5.01	5
6	4.76	8	4.55	9	4.63	7	5.01	8
7	4.66	2	4.41	1	4.61	8	4.82	2
8	4.24	1	4.35	8	4.18	9	4.61	1
9	4.23	4	4.28	7	4.11	1	4.56	4
Lowest	4.23	9	4.03	4	4.1	4	4.3	9
Aver.	4.8		4.69		4.76		5.01	

Overall Assessments

Because it aimed to produce an overall assessment drawing on the input of a range of different teachers, this project considered different ways of producing the overall results, and it was determined that a decision rule rather than an average would be the most suitable way of determining an overall result. The aim of the overall algorithm was to use a range of score points to represent a region on the scale so that the region reported could represent the degree of agreement between the teachers and the degree of precision or certainty of the assessment.

It was also important that the algorithm seemed common sense to teachers, so it was not calculated numerically but developed intuitively on a case by case basis. The algorithm needed to make sense to Overall Assessors and they had to be able to confidently present the outcomes to participating teachers.

The aim of this algorithm was to calculate an initial result (if there was no more than two points difference between assessments) for review by the Overall Assessor. Experience with the kind of data produced by this trial soon made it clear that there was a significant amount of agreement between teachers, and that differences in individual perception of the mark range was often the cause of discrepant ranges. The larger the number of teachers assessing an individual student the more likely there would be a difference of more than two between some teachers, but when there were large numbers it often seemed that one teacher was a clear outlier, and that, on the basis of the degree of agreement between the other teachers, it was possible to discount such outliers. This experience with the data meant that it was possible to see that some differences of 3 points or ranges of 4 score points, that would usually be thought of as discrepant, could be algorithmically reduced to differences of two and less. Accordingly, the algorithm was developed to deal with a difference of 3 and reduce them to differences of two or less by discarding outliers.

For instance, a range of teachers scores of 1, 2 and 3 would be presented as 1-3 by the algorithm, and a range of 1, 3 and 3 would be presented as 2-3 (moving the 1 to 2). These are quite simple cases, but the patterns of scores can become complicated. A range of teachers' scores of 3, 3, 4, 5, and 6 is usually thought of as discrepant in that it involves a difference of more than two between the score of 3 and 6. But this range of scores gives a plausible region for the performance of such a student on this Key Competency, and the algorithm would suggest a range of 3-5 (discounting the 6) for this range of marks.

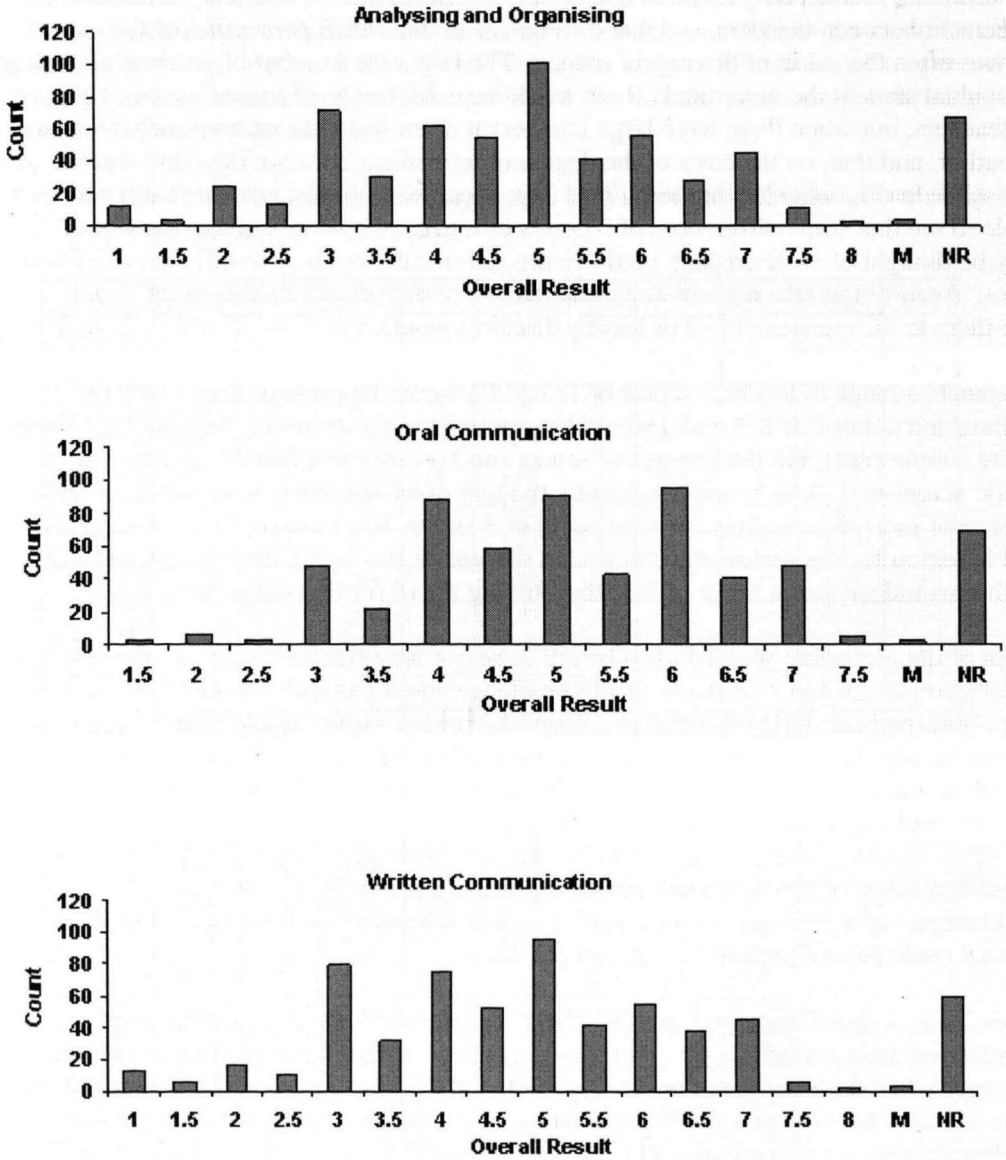
The aim of the procedure was that the Overall Assessor would review such an outcome and either confirm or change it on the basis of knowledge about particular teachers or the student. It turned out during the trial that Overall Assessors were happy with the kinds of decisions the algorithm suggested, and it seems to be a good basis for the initial processing of overall results. As reported below in the NR column of Figure 2, the algorithm could determine results in about 90% of cases in this trial.

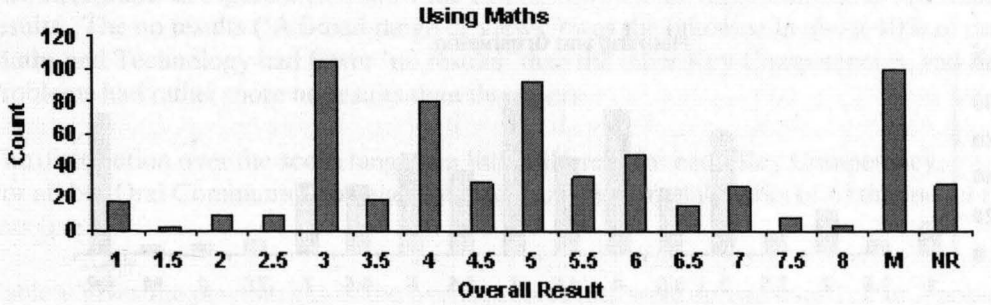
The decision rules for the algorithm are presented in Appendix 8.
Some examples of initial assessments and the overall assessments (the raw data for individual students) are presented in Appendix 9.

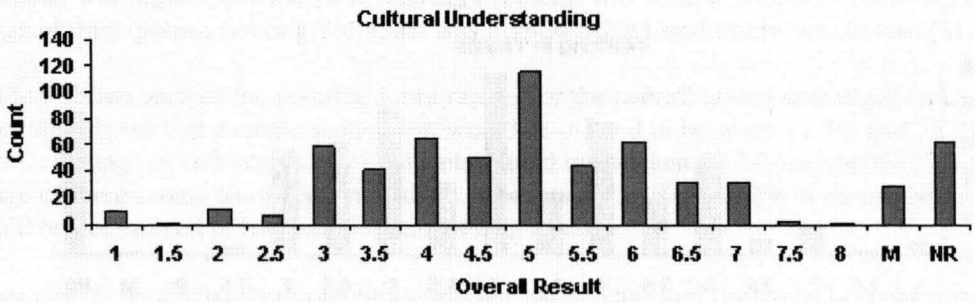
Another issue of significance in assessing the usefulness of these data is the amount of discrimination among students when the assessments of different teachers are combined, particularly when the assessment procedure trialled in this project proposed reporting a range of up to 3 points rather than a single point on the scale. As the distribution of such ranges is difficult to review, an average overall result on a 15 point scale has been calculated and is presented in Figure 2. (In these tables the column M presents percentages where there was only one assessment, and column NR presents the percentage where no result could be determined by the algorithm.)

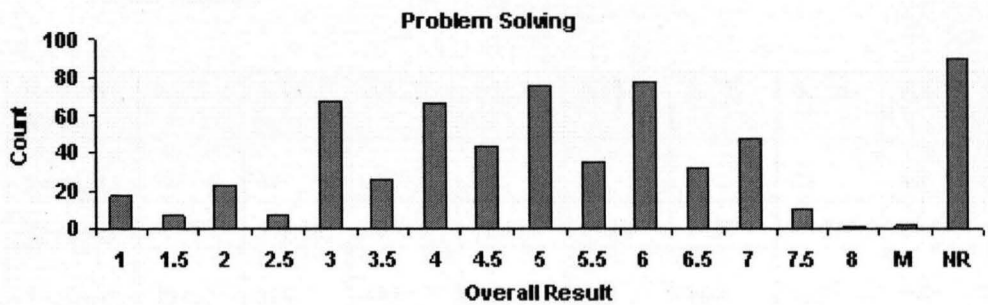
The histograms in general show a distribution of the students between 3 and 7 marks, and the number of students on each point is comparatively even. These histograms suggest that the synthesising of the scores of different teachers did not result in a significant regression to the mean. They suggest that the overall results averaging the assessments of different teachers continue to differentiate among the students.

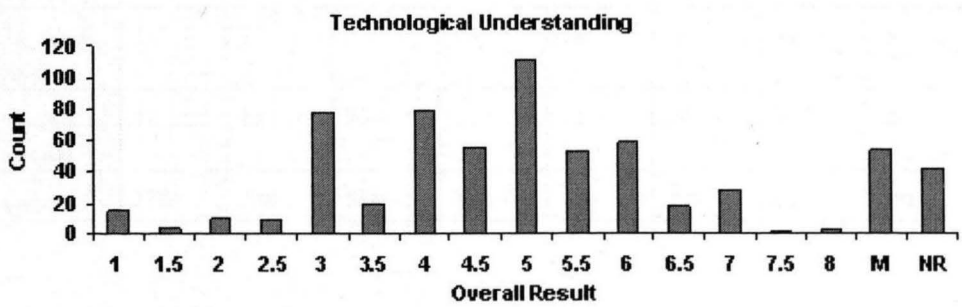
Figure 2 Histograms of overall averages for each Key Competency

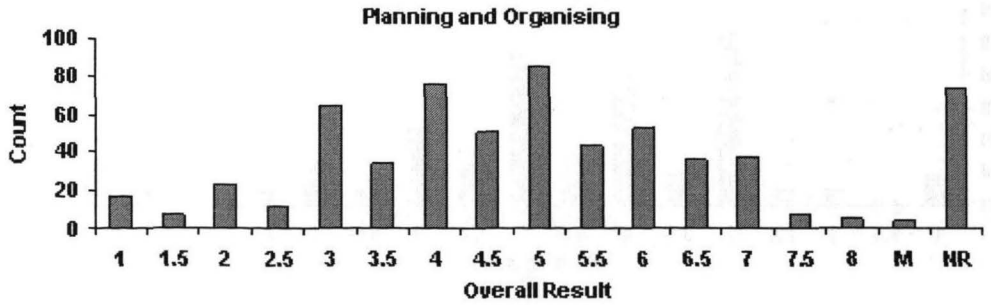


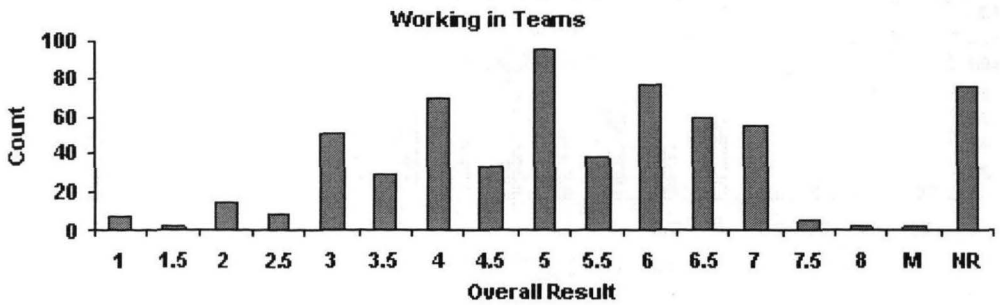

Using Maths
Count
120
100
80
60
40
20
0
1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 M NR
Overall Result


Cultural Understanding
Count
140
120
100
80
60
40
20
0
1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 M NR
Overall Result


Problem Solving
Count
100
80
60
40
20
0
1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 M NR
Overall Result


Technological Understanding
Count
120
100
80
60
40
20
0
1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 M NR
Overall Result


Planning and Organising
Count
100
80
60
40
20
0
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
M
NR
Overall Result


Working in Teams
Count
100
80
60
40
20
0
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
M
NR
Overall Result

The histograms in Figure 2 also show the extent to which the algorithm could determine results. The no results ('A broad range of views') was the outcome in about 10% of cases. Maths and Technology had fewer 'no results' than the other Key Competencies, and Solving Problems had rather more no results than the others.

The distribution over the score range is a little different for each Key Competency. For all but Oral Communication and Solving Problems (modal marks of 6) the modal mark was five.

Table 9 gives the percentages of the overall scores that were spread over 1, 2 or 3 points. Maths had the highest overall results of 1 point (25.9%), and Written Communication (23.7%) and Solving Problems (23.2%) were also among the higher. Team Work had the lowest number of overall results of 1 score (17.7%). For an overall range of two points, Planning was highest (44.1%), and Solving Problems was lowest (39.1%). For an overall range of three points, Solving Problems was highest (42%), and Maths was lowest (31.3%).

Table 9 shows each of the possible score ranges for the overall assessment algorithm. The table shows that a single score point would be offered in between 17.7% and 25.9% of cases. A range of two scores would be determined in between 39.1% and 44.9% of cases. A range of three scores would be determined in between 31.4% and 42% of cases. No result could be determined in between 5% and 14% of cases.

In general is seems that both individual teacher judgements and the overall results spread and discriminated among the students.

Table 9 Percentage of results over I, 2 or 3 points from the overall algorithm for all Schools

Report Range	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 point	93	86	105	71	84	77	87	84	76
Percent	21.281	19.816	23.756	25.912	21.106	18.826	23.262	19.535	17.716
2 point	184	178	185	117	179	160	156	190	176
Percent	42.105	41.014	41.855	42.701	44.975	39.120	41.711	44.186	41.026
3 point	160	170	152	86	135	172	131	156	177
Percent	36.613	39.171	34.389	31.387	33.920	42.054	35.027	36.279	41.259
% Not assess	2	2	3	47	15	3	31	3	3
% No result	11	11	10	5	11	14	6	12	11
cases	558	558	566	500	539	519	535	551	551

Variations between Teachers: the leniency, harshness and consistency of the assessments

The mean for each participating teacher on each Key Competency is presented in Appendix 6, and these teacher summary statistics show some substantial variations between teachers in the way they use the score range. It seems clear that some teachers are substantially more lenient and some are harsher than others.

Some of this variation can be seen by comparing the highest and lowest averages of different teachers in a school. In school 4, for instance, teacher 5 had the lowest average in 8 of 9 Key Competencies. Teacher 6 in the same school had the highest average in 6 or 7 Key Competencies. In school 10 teacher 8 had the lowest average in 5 of 9 Key Competencies. Teacher 11 had the highest average in 5 of 9 Key Competencies.

With comparatively small numbers, it could be these teachers were assessing particularly strong or weak groups of students, and so it is difficult to draw firm conclusions from the means about the leniency or harshness of individual teachers.

An examination of the individual teachers means in Appendix 6 shows substantial differences between individual teachers. In school 1, for instance, the following differences between individual teachers can be seen.

Table 10 The highest and lowest teacher means in school 1.

School One	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
Lower Mean	3.4	3.8	3.1	2.7	3.0	2.5	3.4	3.4	3.5
Higher Mean	5	5.6	5.2	4.4	5.4	5	5.1	4.9	5.3

Similar differences in the mean of scores individual teachers give Key Competencies can be seen in all schools, and in some cases the means suggest there are consistent differences in the way certain individual teachers used the score range.

The extent to which it is the perceptions of the teacher that account for variation in the assessment of individual students can be examined by the comparison of other teachers' assessments of the same students. These differences have been analysed in a fine-grained fashion in the data presented in Appendix 7.

This analysis compares the assessments made by each teacher with every other teacher, taking into account the estimated ability of the students, in a way that is not recognised by a simple comparison of means for each teacher. The analysis compares the expected score for each student calculated on the basis of the scores given by other teachers with the actual score given by a teacher. The difference between the expected and the observed scores are averaged to give a figure for the appropriate adjustment of the range of scores given by a teacher.

In some cases this calculation proposes substantial adjustments up and down for individual teachers' scores. In school 1, for instance, this analysis proposes that teacher 6 on KC1 is particularly harsh and should be adjusted up by 1.4 score points. In the same school teacher 9 on KC1 is lenient and should be adjusted down by .79 of a score point.

Although such changes have not been made in the data analysed here, this method of adjustment could be used to moderate the distributions of individual teachers assessments to make them conform to the pattern of results produced by their colleagues. Such an adjustment would be a means of improving the consistency of teachers' judgements or of narrowing the range of overall results without time consuming consultations between teachers.

The tables in Appendix 7 also present a calculation of the extent to which individual teachers are consistent in the pattern of their scores in relation to other teachers. In school 1, for instance, this calculation shows that Teacher 9 on KC1 is very consistent with other teachers (.17) and Teacher 6 is comparatively inconsistent with other teachers (.68). Analysis of this kind could be used as feedback to individual teachers to help them review and possibly adjust their sense of a Key Competency and their assessment of students in the future.

Reliability of the assessments

The second basic issue in this trial is the extent to which teachers agree in their assessment of individual students. This issue is important as validation of any Key Competencies assessment in which different teacher judgements are to be brought together in an overall school level. Merely averaging any set of results is not appropriate. The crucial issue is what amount of agreement gives a plausible set of marks that can be used to determine an overall result. The more teachers assessing an individual the greater the likelihood that the assessors will vary. But, on the other hand, the more assessments that are available of an individual, the more plausible particular outcomes seem for that individual.

A simple method of attempting to understand the extent to which individual teachers agree in the way they assess individual students is to make pairwise comparisons of every instance that different teachers assess the same student. This comparison will give a percentage of times teachers give the same score, an adjacent score, or a difference of 2, 3, 4, 5, 6 or 7 score points. (For Analysing, for instance, this involves the comparison of the marks of 2177 pairs of teachers.) Table 11 shows the percentage of different levels of agreement in all pairwise comparisons of teachers assessing the same student. Row 0 covers cases where pairs of teachers gave the same grade. Row 1 includes cases where pairs of teachers gave adjacent scores (1/2, 2/3, 3/4 etc.). Row 2 includes cases where pairs of teachers gave a range of 3 scores (1/3, 2/4, 3/5 etc.). And so with rows 3 through to 7.

To understand the significance of these figures, they can be usefully compared with the kinds of percentages that are expected in the external test assessments. In the assessment of a single piece of writing produced in test conditions assessed on specified criteria with described levels of performance over a 10 point scale, it would be expected that trained markers would give the same score in 25-30% of cases. It would be expected that they would differ by 1 in 35- 40% of cases, differ by 2 in 20-25% of cases, and differ by 3 in 5-15% of cases. It is usual for differences of 3 to be judged as discrepancies in external marking and for such differences to be subject to remarking. Typical figures for a 6 point scales are no difference in 35-45% of cases, 1 point difference 40-50% of cases and a difference of 2 in 10-15% of cases.

It was initially proposed during the inception of this project that a difference of 2 (or a range of 3) would be an acceptable degree of agreement to determine an overall result. It was also proposed that 70% of cases within a difference of 2 would represent a substantial degree of agreement on which to construct overall school levels.

Table 11 Assessment Difference Percentages for all Schools

DIFF	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
0	31.328	28.785	32.047	29.630	28.778	27.123	29.890	27.981	27.167
	682	616	681	168	457	578	298	596	561
1	33.119	33.832	33.365	33.157	34.194	32.379	30.090	34.038	33.366
	721	724	709	188	543	690	300	725	689
2	25.540	27.196	24.424	27.513	26.322	25.997	29.990	26.150	27.748
	556	582	519	156	418	554	299	557	573
3	6.890	6.682	7.106	7.937	6.486	9.151	6.820	8.169	7.167
	150	143	151	45	103	195	68	174	148
4	2.572	3.318	2.682	1.587	3.401	4.176	2.608	3.192	3.632
	56	71	57	9	54	89	26	68	75
5	0.413	0.187	0.235	0.176	0.504	0.939	0.201	0.423	0.678
	9	4	5	1	8	20	2	9	14
6	0.138		0.141		0.315	0.235	0.401	0.047	0.242
	3		3		5	5	4	1	5

Table 11 shows some variation for different Key Competencies, but there is a general pattern of substantial agreement on the eight point scale. The critical issue for assessing the agreement in Table 11 is the percentage of comparisons where the difference is more than 2 (usually thought of as a discrepancy), and in general it ranges between 8% and 12% for the different Key Competencies. This means that in around 90% of cases there was a difference of two or less between any two teachers assessing the same student.

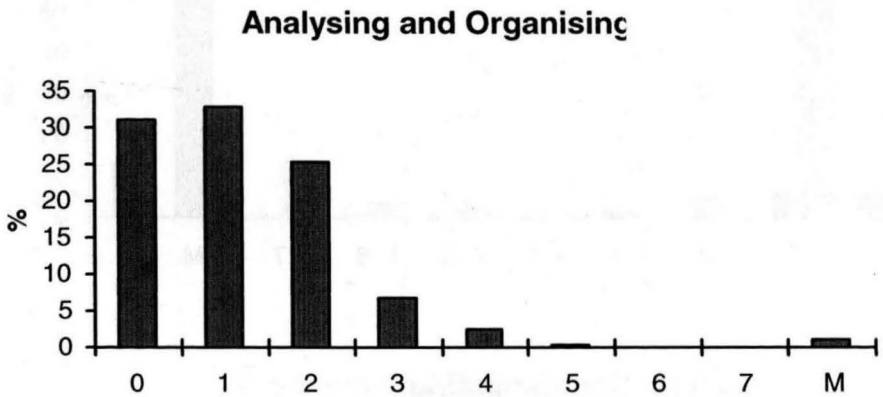
In this trial, between:

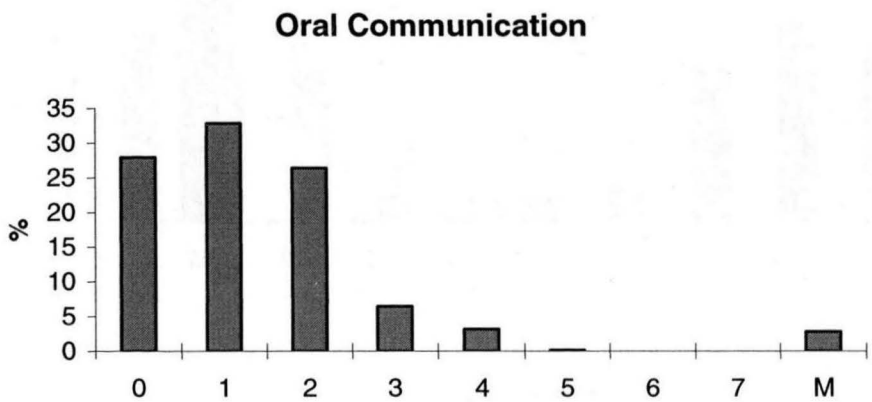
- 27% and 32% of cases pairs of teachers gave the same score;
- 30% and 34% of cases pairs of teachers gave adjacent scores;
- 24% and 29% of cases pairs of teachers gave scores that differed by two points;
- 10% and 14% of cases pairs of teachers gave scores that differed by three points;
- a total of 59% and 70% of cases pairs of teachers gave the same or adjacent scores; and
- a total of 86% to 90% cases pairs of teachers were within the non-discrepant range.

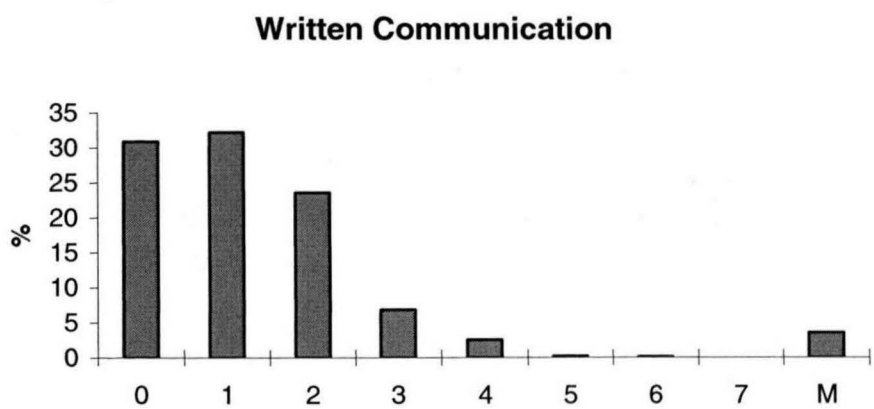
Figure 3 on pages 50 to 52 shows histograms of percentages of agreement between all teachers assessing the same student.

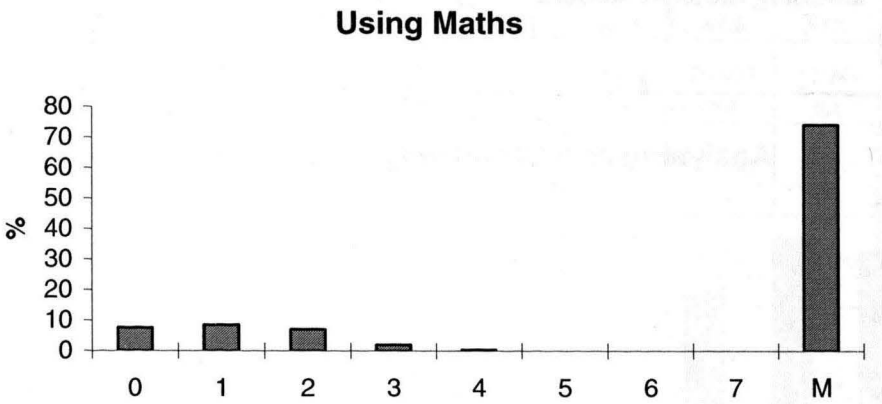
Even allowing for the difference between an 8 and 10 point scale, the degree of agreement between individual teachers in this trial is comparable with that obtained in assessments by trained external markers assessing the same pieces of work in supervised marking procedures. There was a high level of agreement between different teachers assessing the same student on the whole range of Key Competencies in this trial - a level of agreement comparable to that attained in external assessments of single pieces of work.

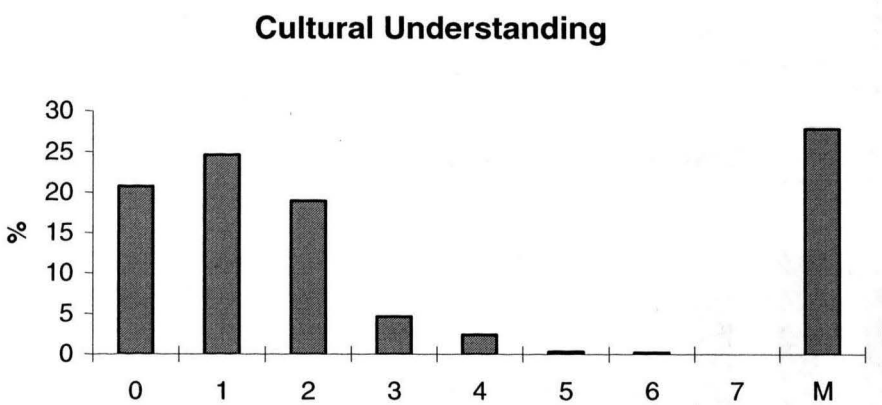
Figure 3 **Histograms of percentages of agreement between all teachers assessing the same student.**

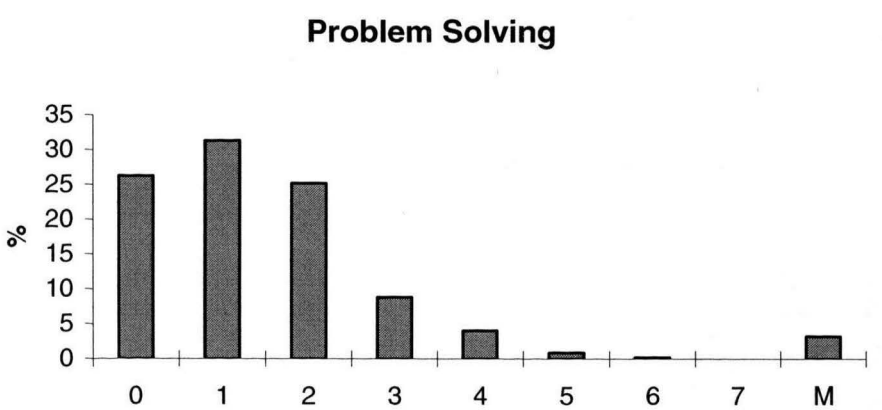


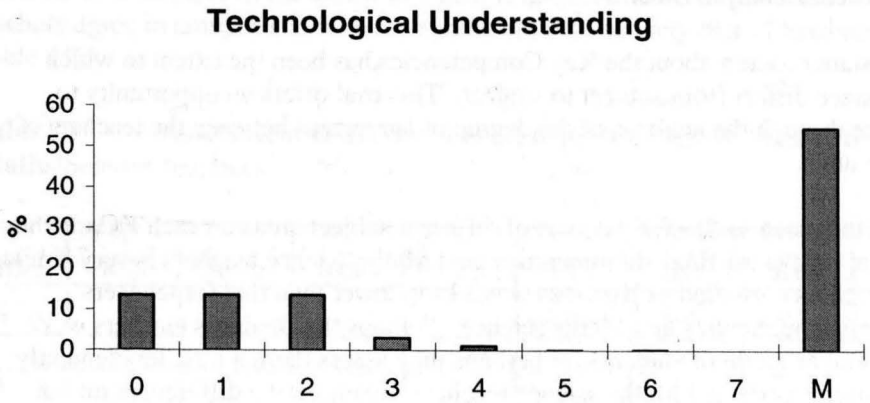


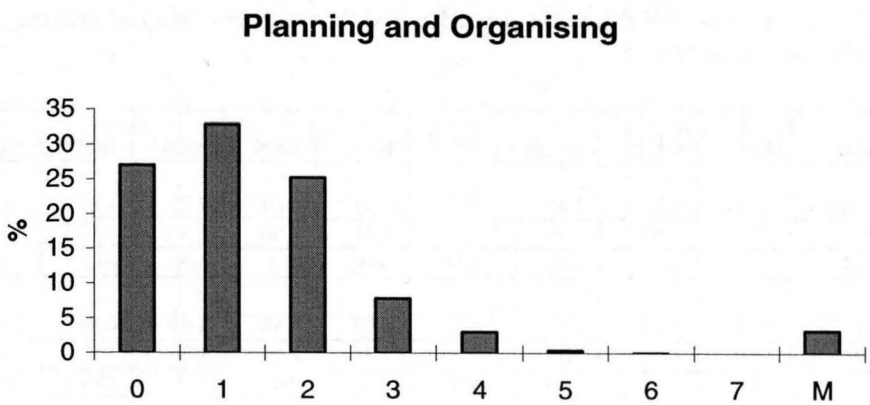


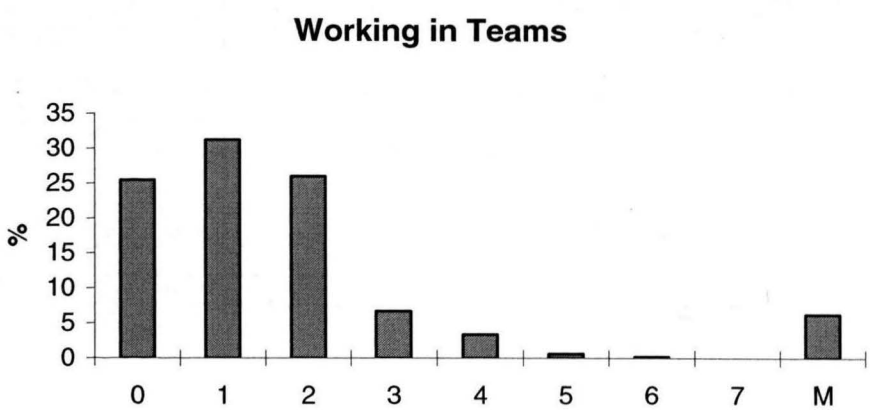

Using Maths
%
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
M


Cultural Understanding
%
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
M


Problem Solving
%
35
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
M


Technological Understanding
%
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
M


Planning and Organising
%
35
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
M


Working in Teams
%
35
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
M

Inter-subject teacher comparisons

An issue of constant concern about the Key Competencies has been the extent to which student performance differs from subject to subject. This trial offers an opportunity to explore this issue through the analysis of the degree of agreement between the teachers of different subject areas.

Table 12 shows the mean scores for teachers of different subject areas on each KC. It shows that the pattern of results for English/humanities and Maths/science teachers is very similar. The mean for teachers classified as Business was a little lower than that for teachers classified as English/humanities and Maths/science. Perhaps the Business teachers work with a slightly weaker group of students, or perhaps they assess them a little less leniently that the English/humanities and Maths/science teachers. Even so, the differences do not seem substantial.

Table 12 Means scores of Subject Groups: English/humanities, Maths/science, Business and Other

ect	Statist	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
Eng/hum	Average	4.96	5.12	4.87	4.47	4.95	4.89	4.79	4.87	5.08
	StDev	1.56	1.50	1.58	1.54	1.51	1.60	1.50	1.62	1.58
	Count	644	641	644	124	643	632	319	642	641
Maths/sc	Average	4.87	5.13	4.95	4.76	4.99	4.89	4.75	4.89	5.18
	StDev	1.60	1.45	1.49	1.56	1.38	1.64	1.39	1.56	1.48
	Count	589	585	574	550	426	586	511	580	584
Business	Average	4.35	4.60	4.36	4.20	4.11	4.46	4.62	4.35	4.64
	StDev	1.69	1.51	1.77	1.65	1.50	1.92	1.64	1.72	1.89
	Count	249	244	245	153	171	250	219	250	230
Other	Average	4.46	4.93	4.52	4.29	4.63	4.71	4.45	4.58	4.85
	StDev	1.65	1.41	1.65	1.46	1.62	1.80	1.58	1.76	1.66
	Count	308	308	307	127	289	305	204	300	298
Total	Average	4.76	5.02	4.77	4.57	4.81	4.80	4.69	4.76	5.01
	StDev	1.62	1.48	1.61	1.57	1.52	1.71	1.50	1.65	1.61
	Count	1790	1778	1770	954	1529	1773	1253	1772	1753

Table 13 is a pairwise comparison of all cases where a student was marked by an English/humanities teacher and a Maths/science teacher.
The table is presented to compare the positive and negative differences (for the English/humanities teachers) which gives an indication of the extent to which one group was harsher or more lenient than the other. The differences are totalled in the last two rows to give a negative and a positive total of difference for English/humanities teachers in comparison with Maths/science teachers. A higher total on the minus side indicates that English/humanities teachers were more lenient than Maths/science teachers, and a higher total on the plus side indicates that English/humanities teachers were less lenient than Maths/science teachers.

The comparison shows a similar pattern in all cases which suggests that there are few systematic differences in the extent to which English/humanities teachers and Maths/science teachers agree in comparison with the agreement between any pair of teachers (as shown in Table 11).

Table 13 Assessment Difference Percentages for English/humanities with Maths/Science teachers

Differ	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
-6					0.211	0.304	0.361		0.149
					1	2	1		1
-5	0.589		0.150		0.211			0.300	0.149
	4		1		1				2
-4	1.031	1.196	1.353		1.899	2.128	2.166	1.652	1.794
	7	8	9		9	14	6	11	12
-3	2.504	4.185	3.459	5.926	2.954	3.647	1.805	3.754	3.737
	17	28	23	8	14	24	5	25	25
-2	10.015	13.004	10.827	14.815	9.072	12.158	13.357	12.312	17.489
	68	87	72	20	43	80	37	82	117
-1	15.317	13.901	12.932	14.074	16.878	14.894	16.245	14.414	13.303
	104	93	86	19	80	98	45	96	89
0	31.959	30.493	34.286	38.519	32.278	28.875	31.047	31.532	26.756
	217	204	228	52	153	190	86	210	179
1	20.619	20.329	18.947	13.333	17.511	18.541	14.801	19.069	21.525
	140	136	126	18	83	122	41	127	144
2	13.108	13.154	14.135	12.593	15.823	13.222	15.884	12.162	11.211
	89	88	94	17	75	87	44	81	75
3	3.535	2.093	2.707		2.532	4.711	3.610	3.604	3.139
	24	14	18		12	31	10	24	21
4	1.325	1.644	1.053	0.741	0.633	1.216	0.722	1.051	0.747
	9	11	7	1	3	8	2	7	5
5						0.304		0.150	
						2		1	
6			0.150						
			1						
- tot	29.3	32.1	28.5	34.7	30.9	32.9	33.7	32.3	36.3
+ tot	38.5	37	36.7	26.5	36.4	37.9	34.9	35.8	36.5

Gender comparisons

There is reason for thinking that the kinds of global impression judgements reported in this study might be more than usually prone to subjective contaminations like differences in the gender of assessors and the students assessed.

Table 14 shows the means for girls and boys on each Key Competency. The mean for girls was higher on all Key Competencies, but the differences between girls and boys varies across the Key Competencies. The difference between girls and boys was highest for Planning and Organising (.77) and Teamwork (.75) and lowest on Maths (.23) and Technology (.2). The difference between the girls and boys on the most verbal Key Competencies., Oral C. (.6) and Written C. (.61), were not the highest, and were not as high as the differences on Solving Problems (.67). These results are comparable to the differences between the performance of girls and boys in other school based assessments.

Table 14 The mean performance for girl and boy students

		KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
Girls	Aver	4.94	5.24	4.99	4.65	4.96	5.05	4.76	5.04	5.29
Total	StDev	1.60	1.43	1.57	1.57	1.48	1.63	1.47	1.59	1.53
	Count	1126	1120	1121	599	992	1122	804	1119	1102
Boys	Aver	4.46	4.64	4.38	4.43	4.52	4.38	4.56	4.27	4.54
Total	StDev	1.62	1.49	1.61	1.58	1.55	1.74	1.54	1.64	1.65
	Count	664	658	649	355	537	651	449	653	651
Differ.		.48	.6	.61	.22	.44	.67	.2	.77	.75

Another issue of interest that can be explored in these data is whether the gender of the teachers seems to relate to the assessments they make.

Table 15 presents the means for teachers of each gender assessing students of both genders. The mean for female teachers was consistently higher than for male teachers assessing all students. Female teachers seemed to rate girls and boys similarly, and female teachers have higher assessments of boys than girls in KC1, KC2b and KC3. Overall it does not seem that there is much difference in the way female and male teachers rate girls and boys.

Table 15 Means for female and male teachers assessing girls, boys and all

	Statist	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
Female	Aver	4.91	5.26	4.86	4.59	4.96	5.09	4.82	5.04	5.31
/Girl	StDev	1.58	1.45	1.57	1.55	1.50	1.60	1.49	1.59	1.54
	Count	682	676	677	371	613	682	472	676	670
Female	Aver	4.98	5.20	5.18	4.76	4.97	4.97	4.67	5.03	5.26
/Boy	StDev	1.62	1.39	1.54	1.59	1.44	1.68	1.43	1.59	1.50
	Count	444	444	444	228	379	440	332	443	432
Male	Aver	4.64	4.81	4.65	4.75	4.82	4.66	4.95	4.55	4.80
/Girl	StDev	1.67	1.48	1.64	1.52	1.59	1.71	1.48	1.71	1.71
	Count	265	266	256	148	200	263	154	262	264
Male	Aver	4.35	4.53	4.20	4.19	4.34	4.19	4.36	4.09	4.37
/Boy	StDev	1.58	1.48	1.56	1.58	1.50	1.74	1.54	1.57	1.59
	Count	399	392	393	207	337	388	295	391	387
F Tot	Aver	4.83	5.13	4.81	4.63	4.92	4.97	4.85	4.90	5.17
	StDev	1.61	1.48	1.59	1.54	1.53	1.64	1.49	1.64	1.61
	Count	947	942	933	519	813	945	626	938	934
M Tot	Aver	4.68	4.89	4.72	4.49	4.67	4.61	4.52	4.59	4.84
	StDev	1.63	1.47	1.63	1.61	1.50	1.75	1.49	1.65	1.61
	Count	843	836	837	435	716	828	627	834	819
All	Aver	4.76	5.02	4.77	4.57	4.81	4.80	4.69	4.76	5.01
	StDev	1.62	1.48	1.61	1.57	1.52	1.71	1.50	1.65	1.61
	Count	1790	1778	1770	954	1529	1773	1253	1772	1753

Table 16 presents the differences between all occasions where female and male teachers assessed the same student. The table shows that the agreement between female and male teachers is much the same as that between English/humanities and Maths/science teachers or between any pair of teachers.

Table 16 Assessment Difference Percentages for Male and Female Teachers

Differ	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
0	30.448	28.978	31.697	30.888	28.507	26.724	27.812	27.399	26.009
	292	275	297	80	191	248	136	257	232
1	32.638	32.244	30.950	30.888	31.343	31.358	28.221	33.902	32.623
	313	306	290	80	210	291	138	318	291
2	25.756	27.713	25.507	25.483	26.866	26.616	34.765	26.333	27.915
	247	263	239	66	180	247	170	247	249
3	7.091	6.428	7.471	10.039	8.060	9.591	6.748	7.889	7.848
	68	61	70	26	54	89	33	74	70
4	3.233	4.426	3.842	2.317	4.478	4.849	2.249	3.518	4.596
	31	42	36	6	30	45	11	33	41
5		0.211	0.213	0.386	0.597	0.754		0.853	0.785
		2	2	1	4	7		8	7
6			0.320		0.149	0.108	0.204	0.107	0.224
	2		3		1	1	1	1	2
Total	959	949	937	259	670	928	489	938	892

Table 17 compares the assessment differences between any and all pairs of teachers with the pairing of male and female teachers and the pairing of English/humanities and Maths/science teachers. This comparison shows a very stable pattern of agreement with a discrepancy rate of no higher than 15% in any case.

Table 17 Comparison of assessment differences for All teachers with Male and Female teachers, and English/humanities and Maths/science teachers

Differ	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
0-1 All	64.4	62.5	65.3	62.7	62.8	59.4	59.8	61.9	60.4
0-2 All	89.9	89.6	89.7	90.2	89.1	85.3	89.7	88	88.1
0-2 M/F	88.7	88.9	88	87	86.6	84.6	90.7	87.5	86.5
0-2 E/M	90.9	90.7	87.1	93.4	91.3	87.5	91.2	89.3	90.2

The agreement reported in the table above is high, and substantially exceeds the 70% of cases within a difference of 2 which was proposed at the inception of this project as a workable basis for creating overall school levels.

Relationship between the Key Competencies

Table 18 below shows the correlations between the scores of each student on all Key Competencies. The correlations vary. Some of them are moderate (.5) and some are quite high (.8). The highest correlations are between Analysing and Written C. at .83 and between Solving Problems and Planning at .83. The lowest correlations are between Written C. and Maths at .5, and Cultural U. and Technology at .53.

This pattern of correlations is quite intelligible, and it suggests a literacy/oracy/humanities factor, a solving problems/planning/teamwork factor, and a science/technology factor.

Table 18 Inter-correlations of Key Competency judgements

	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
KC1		.67	.83	.78	.71	.77	.64	.75	.65
KC2a			.70	.59	.67	.65	.55	.65	.69
KC2b				.74	.70	.75	.62	.74	.65
KC3					.52	.72	.67	.67	.57
KC4						.69	.53	.70	.65
KC5							.69	.83	.72
KC6								.68	.58
KC7									.73
KC8									

Table 19 show the correlations between Maths/science and Humanities teachers, Maths/science and Business teachers, Humanities and Business teachers, Female and Male teachers, and any pair of teachers.

Maths/science and English/humanities teachers agree most on Maths (.57) and least on Technology (.30). Maths/science and Business teachers agree most on Written C. (.60) and Maths (.58) and least on Oral C. (.33) and Technology (.37). English/humanities and Business agree most on Analysing (.38) and least on Technology (.15).
Females and Male teachers agree most on Maths (.49) and least on Cultural U. (.3).

Table 19 Inter-correlations of Math/science and English/humanities, Maths/sciences and Business, Female and Male teachers, and any pair of teachers

	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
M/E	.49	.43	.48	.57	.41	.44	.30	.47	.40
M/B	.48	.33	.60	.58	.46	.40	.37	.51	.51
E/B	.38	.27	.34	.21	.43	.35	.15	.43	.43
F/M	.44	.34	.44	.49	.30	.39	.35	.42	.37
Pairs	.46	.37	.47	.46	.35	.40	.32	.44	.39

In conclusion

The analysis of the assessment data from the trial showed that:

- participating teachers were able to discriminate among their students and spread them across the mark range using the Key Competencies framework;

- significant numbers of teachers did not assess Using Mathematical Ideas and Techniques, Using Technology and Using Cultural Understandings, but almost all participants were able to assess the other Key Competencies;

- there were some systematic variations in the way the teachers used the mark range, (there were some comparatively harsh and lenient teachers), and there were differences in the consistency of individual teachers' assessments with their colleagues (there were some relatively consistent and some inconsistent teachers);

- there was quite a high level of agreement between different teachers assessing the same student on the whole range of Key Competencies (a level of agreement comparable to that attained in external assessments of single pieces of work on similar mark ranges);

- the level of agreement between teachers about individual students in this trial makes it realistic to construct overall and cross-curricular assessments of Key Competencies;

- the data offered little indication that the perspectives of teachers in different subject groupings differ substantially or systematically from teachers in other subject groupings;

- the data offered little indication that the perspectives of female and male teachers differ substantially or systematically in the way they assess either girls or boys; and

- there was a fairly consistent pattern in the way individual schools assessed their students in comparison with other schools.

In a number of respects the results of this trial are surprising and remarkable. They suggest that such assessments are psychometrically sound, and that they carry useful information. The teachers within a school as a group seem to discriminate consistently among their students. And although these data do not address the issues of the reliability of comparing assessments from different schools, they do show that the assessments within schools have a significant degree of reliability.

This trial suggests that the Key Competencies can be operationalised in a coherent fashion, and they can be seen and enacted as generalised notions that hold their meaning across different subject areas or domains. This trial also suggests that the Key Competencies can be usefully assessed at provider level without a large investment of time or resources.

Part 5 Reactions of Employers to a Sample Report Based on the Mayer Key Competencies

This part of the project aimed to obtain responses from employers to a school report based on teachers' assessments of students' performance on the Key Competencies. A complete report of this phase of the study is given in Appendix 10.

The format of report shown on pages 61 and 64 was circulated for comment.

It seemed important to obtain a national perspective on the reactions of employers to this kind of reporting because the scheme, when implemented, would entail employers across the country being able to make effective use of the information available in this kind of report. The survey also sought opinions from both large and small businesses.

A questionnaire survey and telephone interview were used - the survey to gather a wide range of responses and the interview to obtain specific and more detailed comments on certain aspects of the sample report - in particular to discuss suggestions for improvement.

Details of the numbers and kinds of business approached and the response rate are discussed in Appendix 10.

Nine main findings emerged from this study.

1. Respondents found the Key Competencies to be a good basis for schools reporting student performance to employers but they require other information as well.

The kinds of additional information required included information about any prior work and in the case of school leavers, of work experience. In some instances assessments of performance in specific subject areas would be required.

2. Respondents found the information provided in the Sample Report to be good, but for most its presentation was too complex and unclear. This is particularly the case with elements 1 and 2. (reporting on the Key Competencies and the highlighting of particular facets of the Key Competencies)

Elements that were difficult to follow included the graphic display and the facets.

3. Levels of performance information is welcomed by almost all respondents. The concerns of those who do not welcome it are focused largely on presentation and lack of clear information about the meaning of what is before them.
Clearer specification of the standards communicated by the scale and clearer communication of the direction of the scale are required.

There was general agreement that the levels of achievement information was to be welcomed, there was a range of views about how much confidence could be placed in it. The kinds of 'soft skills' that employers look for in potential employees agree very closely with the kinds of attributes covered by the Key Competencies. Respondents look for other information, ranging from academic results to information about work experience and career interests.

Figure 4 **The Key Competencies sample report for employer consultation**

School Reference and Key Competencies Portfolio

Kingston High School

Smith Road, Kingston

Phone: 9677-6788 Fax: 9677-9000

Issued for

Jill Johnson

on 31st December 1996

The Key Competencies are a mixture of the knowledge, skills and understandings that all young people need to enable them to participate effectively in the workforce. The Key Competencies, and the Facets of each, are outlined on the back page of this Report. This School Reference and Key Competencies Portfolio Report has 2 sections and 6 Parts. Section 1 [page 2] is produced by the school, and Section 2 [page 3] is produced by the student.

Section 1

Part 1 shows, for each Key Competency, one or up to three sub-levels of performance determined on the basis of the judgements of a group of teachers. A narrow band means teachers agreed closely;a broader band indicates a wider range of teacher judgements.
Part 2 identifies specific Facets of Key Competencies that are typical features of, or particular strengths of, the student's performance using a Facet descriptor or a specific comment.
Part 3 is a statement made by the school on the basis of a group of teacher judgements and comments about the performance of the student in a range of activities. Further detail about these activities may be found in Section 2

Section 2

Part 4 is the student's statement about her or his participation in various school activities and her or his performance on the Key Competencies.
Part 5 is a statement by the student about specific skills that she or he is able to demonstrate.
Part 6 is an index of items offered for inspection in the student's Portfolio

1

Issued for : Jill Johnson **By :** Kingston High School

Date: 31st December 1996

1	Key Competencies	Not Yet Level 1	Level 1	Level 2	Level 3
KC 1	Analysing and Organising				
KC 2	Communicating				
KC 3	Mathematics				
KC 4	Cultural Understanding				
KC 5	Problem Solving				
KC 6	Technology				
KC 7	Planning				
KC 8	Teamwork				

2	School Facet Highlighting or Specific Comments		
	This student has a particular strength in (KC1-4) /or typically (KC5-8) (see back page)	KC No	Facet
KC 1	collecting information that is relevant		
KC 2	using communication skills to interact with other people	KC 2	C
KC 3	understanding mathematical principles and procedures	KC 3	B
KC 4	developing a coherent position on social issues.	KC 4	B
KC 5	shows focus and persistence	KC 5	A
KC 6	is thorough and persistent in dealing with technology.	KC 6	B
KC 7	is flexible and innovative.		
KC 8	leads or facilitates group processes.	KC 8	B

3	School Statement

Jill was a very positive and constructive member of our community during her four years at the school.

Jill has taken part in a number of school bodies. She was a member of the SRC for two years and she was part of the Year 11 and 12 Dance committees. Jill also played in the school netball team and participated in athletics and swimming competitions within the school.

Jill has shown a particular interest in textiles and sewing over a number of years. She organised her school program around the intention to join the fashion industry, and she has pursued this aim systematically. Jill organised for herself two successful periods of extended work experience with local dressmakers.

Issued for : Jill Johnson **By :** Kingston High School

Date: 31st December 1996

4	**Student Statement**

I believe that I have made good use of my time at school, and that I have got a lot out of it. I was involved in sport and other activities. I took a leading role in the Dance Committee. I was also on the SRC.
The school helped me develop a focussed approach to my school work and to set myself some goals. I improved my reading and writing a lot in Years 11 and 12. I have always liked sewing and it is work that I would really like to do. It is difficult but it is satisfying. I organised good work experience with local tailors.
I also believe that I can get on well with people and deal with them, even when they are impatient and rude.

5	**Student Statement of Specific Skills**
1	drafting patterns and reading and constructing advanced patterns
2	use of computerised sewing machine and overlocker
3	hand finishing of garments to a good level
4	use of Word for Windows at a competent level
5	touch typing at thirty words per minute
6	use of Excel spread sheets for accounts
7	working a supermarket checkout and managing a till
9	first aid at certificate level

6	**Student Index of Portfolio Items**
1	Ball dress photo display
2	Tailored suit photo display
3	Embroidery samples
4	Essay on Euthanasia from Year 12 English
5	Year 12 Accounting assignment on small business management
7	Work experience references
8	Reference from supermarket supervisor
9	First aid certificate

There was strong support for the reporting of levels of performance and, with two exceptions, a preparedness to accept the judgements provided as valid and useful.

4. The concept of facets and the descriptors provided were widely welcomed by respondents and the concept of distinctive comments was also strongly supported. However, important improvements are required.

No respondent suggested the need to reconceive the major categories represented in the Key Competencies themselves and only a few comments were made suggesting modification of the way the Key Competencies were broken down into facets.

Discomfort was expressed, however, about the way some of this material was presented. For example, two respondents misunderstood the A/B/C code thinking that it indicated a hierarchy of levels of performance (as can be seen in Part 3, a few teachers were also confused by this).

One detailed comment raised the issue of the conceptual links between the summary, the facets and the school statement. At the moment, the format does not clearly link these elements.

In summary three improvements need to be made.

1. Render the information in as clear and as graphic a fashion as possible. For example, symbols rather than numeric codes could be employed; unique comments could be highlighted to distinguish them from standard descriptors.

2. Provide explications (they would need to be brief) of the Key Competency scale, the facets (rather than having to turn a page) and the relationship between each of the elements of the school's assessment. These requirements may present significant but not insuperable challenges to layout-design and copywriting.

3. Render the relationship between summary global judgements, the facet descriptors/distinctive comments and the continuous prose of the school statement transparent to the user.

5. The school statement is very strongly supported and should be seen to explicate and elaborate the summary information which precedes the Sample Report.

This element is the one most closely resembling the type of reference that schools currently issue. Approximately 80 per cent of respondents thought that this element would be useful and informative.

6. The monitoring of assessments produced by schools is seen to be highly desirable by the great majority of respondents, although some were so strongly in support of the Sample Format that they felt they could not be more positive than they were.

7. There is clear support for the inclusion of a student statement but concerns voiced about how it can be authenticated and therefore equitably regarded. The issue is to be resolved as to whether the school does this or whether the student is regarded as a free agent in assembling this and other material.

The statement by students was favourably regarded because it provides an opportunity to consider the applicant in their own voice. Employers believed that this would give them a sense of the applicant's personality. But concern was expressed about the extent to which the text produced would be the unaided work of the student.

8. There is general support for the inclusion of Specific Skills in the list provided by the student. This list will be more useful if the information included in it is precise and certificated or authenticated. It will be in the student's interest to ensure that this is the case.

The list of specific skills was warmly welcomed but some concern was expressed about the relevance of the list to specific jobs and job seeking processes. This concern could be addressed by preparing different versions tailored to specific job-seeking initiatives. It will be more difficult to ensure that skill levels indicated are accurate and comparable.

9. There is general support for the inclusion of a list of portfolio items especially if the selection is targeted to a particular job application and related to the Key Competencies.

Respondents stressed that the recentness of items and the relevance to a job application would be important.

As mentioned above there is a much fuller discussion of data in the report (Appendix 10). The strength of these data rest to a large extent on the expertise of the respondents - some from senior human relations management positions in major Australian companies. The response rate is low for various reasons outlined in the full report. Reasons are given which indicate that the responses cannot be taken to represent 'employer views', but the findings seem to be consistent with those of other studies.

Bibliography

Mayer, E. (Chair) (1992): **Key Competencies.** Report of the committee to advise the AEC and MOVEET on employment-related Key Competencies for postcompulsory education and training. Australian Government Publishing Service, Canberra.

McCurry, D. (1996) Approaches to Assessing the Key Competencies Camberwell, Vic.: ACER

National Industry Education Forum (1996) Business, Industry, Key Competencies & Portfolios

Appendix 1 An extract from the report on the assessment of Key Competencies to DEETYA by McCurry (1996)

APPROACHES TO ASSESSING THE KEY COMPETENCIES

Summary

This report is based on a review of the Key Competencies projects supported by the Department of Employment, Education and Training. It has been developed from discussions at various meetings of officers from the Key Competency projects and officers of the various curriculum and assessment boards in the states and territories.

The Key Competencies projects in the states and territories have raised many questions and concerns. Doubts have been raised about the validity, reliability and feasibility of assessing the Key Competencies. The aim of this work is to review various causes of concern and to see how various difficulties might be obviated in developing a national Key Competencies assessment.

The aims of this report are to:

- review the suggestions in the Mayer Report about the assessment of Key Competencies;
- review the projects in different states and territories dealing with assessment of Key Competencies;
- review the issues to be taken into account in developing a Key Competencies assessment; and
- present and analyse a small number of models for the assessment of Key Competencies.

As a result of this work, the following positions were developed.

1 Validity and reliability are relative concepts. Reviewing the validity and reliability of Key Competencies assessment regimes depends on the purpose and function of the assessment.

2 Clarification of the purpose and function is the crucial first step in developing a Key Competencies assessment. Evaluating the validity and reliability of a Key Competencies assessment depends on the purpose and function of the assessment.

3 The essential feature of the Key Competencies is that they are nominated as the abilities sought in employees rather than a record of achievements in the school curriculum. The Key Competencies have a significantly different emphasis and scope from most educational assessments. The Key Competencies are less cognitive, more performance oriented and more attitudinal than most educational assessments. These differences mean that while the Key Competencies can readily be seen as part of the curriculum, Key Competencies assessments cannot be readily amalgamated with subject assessments without diluting them both.

4 The Key Competencies are not especially problematic constructs, and they are amenable to credible assessment. They could be assessed with the hard-edged reliability required by some purposes or the soft-edged reliability suitable for other purposes.

5 The proposals for assessment outlined in the Mayer Report would not attain the kind of hard-edged reliability that could support system-level and national monitoring for public accountability or selection into further education.

6 It is unrealistic to envisage a Key Competencies assessment system as competing with or replacing subject-based grades in the school system. It would be desirable for a Key Competencies assessment to offer different information to that offered by subject-based grades. Where system wide courses do not exist as the basis for reporting achievement, the Key Competencies could offer a framework for reporting levels of performance.

7 Regimes that adapt subject-based assessments to be Key Competencies assessments may either confound the subject assessments or they may be limited in what they can assess. Such adaptations would entail significant changes to current curriculum and assessment practices in most states and territories.

8 It is doubtful whether a feasible or cost effective system of standardised assessment tasks for assessing Key Competencies (and including all providers) could be developed. Similarly, it is doubtful whether a system of standardised assessment tasks could be developed so as to produce data that can be used for hard-edged comparisons within a system or across the country. For such a purpose, studies of representative samples of students could be mounted that would support hard-edged comparisons within systems and across the country.

9 The kind of contextualised, informal and provider-based assessment envisaged in the Mayer Report can attain a kind of soft-edged reliability that could meaningfully contribute to the reporting necessary for transition from school to work for students and the process of staff selection by employers.

10 The practical option for the assessment of Key Competencies is a system where judgements are made by accredited providers according to guidelines administered by a central agency. Initial Key Competencies assessments would be made by class teachers on the basis of what they observe in and out of class. These initial assessments would be synthesised into an overall level score for each Competency. The scores would be reported on a certificate under the auspices of and produced by the provider. The level scores awarded by each provider would be reviewed and monitored by a central agency.

11 The distinctive characteristic of the Key Competencies is the way they can facilitate transition from school to work for students and the selection of employees. Employers want more information in selecting employees than is offered by subject grades. There is reason for thinking that employers will want Key Competency levels and more than levels. Graded information going from summary levels and standardised descriptors through to unique descriptions and even student portfolios should be the aim of a Key Competencies assessment.

Table 4A PROPOSAL FOR A SUMMATIVE PROVIDER-BASED ASSESSMENT OF KEY COMPETENCIES

	An Informal, Contextualised and Provider-based Assessment of Key Competencies
Scope	• covering all KCs • based on as many subject areas as are meaningful for each KC
Contexts	• primarily based on subjects and other in-school activities • possible inclusion of out of school activities if they are supported by plausible evidence or can be readily attested • possible inclusion of standardised descriptors and individual descriptions of performance
Auspices	• reporting under the auspices of the provider • accreditation by central body on the basis of specified procedures • provider quality assurance based on system guidelines • monitoring of results by central body
Mode	• global inferences about generic skills • informal judgements contextualised within classwork and other activities • generic skills distinguished from the subject or context specific skills
Method	• global impression judgements made by subject teachers • synthesis of different results into an overall level • publication of overall results to contributing teachers and review of results if necessary
Reporting	• reporting of levels for each KC • Mayer levels used as a base, and adapted into a seven point scale - a single point or up to three points can be reported • optional reporting of standardised and unique descriptions
Monitoring	• submission of levels to central agency • review of levels in relation to results for the system as a whole and in comparison with subject-based assessments

A Proposal for a Provider Based Key Competencies Assessment

The following proposal is developed on the reasoning presented in this report. It envisages a provider-based assessment established within ACACA Board guidelines and with results monitored by the Boards.

A PROPOSAL FOR A PROVIDER-BASED ASSESSMENT OF KEY COMPETENCIES

This Key Competency assessment system would be based on judgements about levels of performance, and with the optional use of standardised descriptors or unique descriptions. These different assessments would be reported under the auspices of the providers.

The providers would be accredited to produce and report the assessments by a Board on the understanding that the provider follows specified procedures and submits candidate results for review by ACACA Boards.

THE PROCEDURE

This assessment would involve all subject teachers registering a score for most Key Competencies, and all Key Competencies if possible.

The assessment would entail global inferences about competencies that are judged to be typical of the student, and that are likely to have wide or general application.

The assessment would be based primarily on participation in classwork and other school activities. Non-school activities could be taken into account if they are well-attested, and the teachers feel they are typical of the student.

There would be three components to the assessment system. The first would be the base requirement for all accredited providers, and there would be two optional components that could be used at the discretion of the providers.

Component 1 Positions or ranges on Key Competencies continua

Teachers will be expected to give each student a position or a range of up to three points on a seven point continuum for each Key Competency. The continua would be generated by dividing each of the three Mayer levels into two points with the addition of a point that is not yet Level 1.

The point or range of points determined by individual teachers will be submitted to a person whose role is to synthesise the subject teachers' views into an overall point or range of points on the continuum for reporting by the school.

The overall decisions would be made on a case by case basis rather than an algorithmic process. The person making the overall decision would take into account the different perspectives of the different teachers and try to strike the most appropriate balance between them.

The assessments from separate teachers and the overall position would be presented for review by the whole staff. Some cases could be subject to question and reconsideration.

If the range of assessments offered by individual teachers was too diverse to be reconciled, the inability to attain consensus could be indicated.

There would be an internal review procedure for students who were dissatisfied with their assessments.

Component 2 Standardised descriptors

This section of the assessment would be based on 30 standardised descriptors for each Key Competency that had been ordered into three levels. There would be a bank of 10 descriptors for each level.

When the teachers were deciding on a score, they would review the set of 10 descriptive statements of general ways in which a student may be judged to be at a certain level. In doing so teachers could record codes for the two descriptors that best reflected their sense of the student's level of competence.

The descriptors selected could be taken into account by the person determining an overall score for a student. The descriptors that are most typical of the student could be automatically collated and produced on a report for the individual.

This form of reporting could be readily produced electronically.

Component 3 Individual descriptions

Teachers may, if they wish, record particular comments or instances as examples of the students competencies when they are selecting a score in Component 1. The comments could be adaptations of the standardised descriptors reviewed for Component 2.

These individualised or unique comments could be used as the basis for a written report of the students' Key Competencies achievements. These comments would be synthesised from the various comments of individual teachers by the person making the overall assessment.

This form of reporting may be produced at the discretion of the provider when requested by the student.

QUALITY ASSURANCE

Although this Key Competencies assessment is made and reported by providers, they would be authorised to issue Key Competencies certificates if they are accredited by a Board.

Accreditation would depend on adherence to the procedures outlined in Component 1 and a satisfactory review of the results presented by the provider from year to year.

Monitoring of results

The range of results produced by a provider under Component 1 would be reviewed each year by a Board. The review would be based on a comparison of the pattern of results given by individual providers with the results for the system as a whole. The monitoring of results would take place on the following grounds.

There should be a fairly consistent pattern of results produced across the system. Providers that awarded markedly higher or lower distributions than were produced in the system as a whole would be asked to explain why their pattern of results were either higher or lower that the system as a whole.

Similarly the pattern of correlations between different competencies would be expected to be fairly consistent across the system, and providers might be asked to explain if the pattern of correlations they produced differed markedly from those produced in the system as a whole.

There would also be a possibility of using subject grades from system-wide assessments to review the results awarded by individual providers in the more academic competencies. Subject results or groups of subject results might be used to rank order schools on the more academic competencies, and comparisons might be made between the relative standing of providers on the basis of their performance on the state-wide subject-based grades and the Key Competencies assessments. Providers that produced results for the more academic Key Competencies that poorly reflected the performance of the same students on subject-based grades would be asked to account for the difference.

A Portfolio Approach

The proposal sketched above is compatible with and could be part of a portfolio approach to assessment.

The aim of the assessment sketched above would be to present general summary information that could be condensed into one page. This single page of summary information could be the opening of a portfolio of specific skills and competencies, other broader evidence, and work samples collected by the student.

Appendix 2 The SBKCLAP Brief

A Project to Explore a School-based Assessment and Reporting of the Key Competencies

a DEET funded undertaken by the

Australian Council for Educational Research

The draft report written for DEET by Doug McCurry of ACER between June 1995 and February 1996 on the assessment of Key Competencies in schools suggested that the most practical and realistic regime for assessing the Key Competencies would involve:

- global impression judgements made by teachers on the basis of normal classwork and other school activities;
- a process where the judgements of class teachers were synthesised into overall levels of performance;
- overall judgements being reported through levels (there could be optional use of standardised descriptors and unique descriptions); and
- results being reported by providers, but providers would need to be accredited by, and the results reported would be reviewed by, a central agency.

The aim to this proposed project is to explore the possibility of basing a system of Key Competencies assessments on global teacher judgements made on the basis of normal classwork and other school activities. This project proposes to:

- develop and trial a system of the kind proposed in the report by McCurry in a number of schools to gather reactions from teachers about the feasibility and the value of the procedures proposed and the assessments produced, and
- to analyse and assess the degree of convergence in the assessments of individuals by different teachers in the assessment trial.

The proposed assessment trial is based on the suggestions in the Mayer Report that Key Competencies assessments can be made in different programs without entailing new or different assessment tasks. The proposed assessments would be global impression judgements made by a range of teachers, and these judgements would be subsequently synthesised into overall judgements that could be reported under the auspices of the school.

The project will involve some focussing and interpreting of the Key Competencies for the purpose of teacher-based judgements in general education programs. It will also involve the elaboration of the three levels outlined in the Mayer Report into a broader scale and the development of a method for recording teacher judgements efficiently and conveniently.

The project will develop protocols for the judgement of the various competencies and a method for making and recording teacher judgements will be designed. The project will also involve the development of a process for synthesising the different judgements of teachers into an overall result, and a format for reporting the overall result for each competency.
It is proposed to elaborate the three levels proposed in the Mayer Report into a seven point scale. It is further proposed to offer sets of descriptors for each of the three Mayer levels of each competency as the basis for the global judgements. It is also proposed that these standardised descriptors be used as a method for reporting. A procedure for gathering and reporting unique comments on individuals will also be trialled.

It is also proposed as part of the project to solicit reactions and opinions from employers about the proposed means of making Key Competencies assessments and the methods of reporting them trialled in this project.

Assessment Implements and procedures developed by the project

The project will involve the:

- refinement and focussing of the Key Competencies as the basis for global teacher judgements in general education programs;
- development of a set of standardised descriptors to elaborate the three levels of performance for each competency;
- development of protocols and support materials for the assessment of the Key Competencies;
- development of a procedure for recording judgements and for synthesising judgements into an overall level for reporting on each Key Competency; and
- a format for reporting the overall Key Competencies judgements, including the reporting of standardised descriptors and unique descriptions.

The trialing procedure

The material developed and the proposed procedures will be trialled in nine schools. The object of this trialing will be to determine whether the practitioners who enact the procedures in the trial feel that:

- the basis for the global judgements is intelligible and workable;
- the procedures for making and reporting judgements are manageable and efficient; and
- the overall results are meaningful and valuable.

The trial will also analyse the degree of agreement or convergence of the assessments made of individual students by different teachers.

The project will seek the cooperation of eight schools in three states to trial the proposed procedures. The project schools will be asked to nominate three class groups or about 100 Year 11 students. All the teachers of those classes will be involved in the trial. The participating teachers will be given a project briefing and limited training in the interpretation and use of the assessment and reporting framework developed by the project. They will also be given written support materials to assist in making the assessment.

A method for recording and reporting judgements will also be offered by the project for the use of the schools. This method may be a prototype computer program for recording the judgements of individual teachers, synthesising different judgements and producing reports electronically. There will also be a hard text version that captures the same information.

The object of the trial will be to see whether and to what extent teachers and schools feel that:

- the basis for the global judgements is intelligible;
- the information can be recorded and reported efficiently; and
- the final reports present information of value.

The reactions of the teachers participating in the procedures will be sought through questionnaires and interviews.

The second object of the trial will be to test the degree of agreement or convergence in the way individual students are assessed by different teachers. The degree of agreement will be taken as an indication of the soundness of the of the trial procedures.

A project timetable

May 1996 assessment and reporting framework will be developed

initial piloting of the assessment framework by the NIEF Portfolio project schools

June participating schools will be sought

convening of a consultative committee for the project

July briefing meetings for participating schools
initial support materials will be presented to schools

September final assessment mechanism and briefing materials will be presented to trial schools

Octoberindividual teacher judgements will be made

consultations with employers about the procedures and the method of reporting assessments

Nov individual judgements will be synthesised and reports will be produced
reactions will be sought from teachers about the framework and the procedures

Dec reactions of participants and employers will be sifted and analysed

results of assessment will be analysed

Feb 1997 meeting of Consultative committee to review findings

April a completed project report will be presented

Appendix 3 SBKCLAP Advisory Group membership

Marie Dixon Department of Employment, Education, Training and Youth Affairs GPO Box 9880 CANBERRA ACT 2601
Anne McLeish Executive Director National Industry Education Forum P O Box 7225 MELBOURNE VIC 3004
Sue Halliday National Industry Education Forum P O Box 7225 MELBOURNE VIC 3004
Chris Reynolds Educational Consultant 5 Wordsworth St St Kilda 3182
Elaine Grossman Australian Institute of Petroleum Level 23 500 Collins Street MELBOURNE VIC 3000
Chris Ryan NSW Department of Training and Education Co-ordination GPO Box 33 SYDNEY NSW 2001
Del Colvin Department of Education Floor 9 Education House Mary Street BRISBANE Q 4002
Tony McKay National Curriculum Services IARTV Mercer House 82 Jolimont Street JOLIMONT VIC 3002
Robert King NIECAP Project Officer 9/48 Dalgety Street ST KILDA VIC 3182
Tony Brown PO Box 18 STRATHFIELDSAYE VIC 3551

Sue Dunbar &
Andrew Fitzsimmons
Bradfield College
192 Pacific Highway
CROWS NEST NSW 2065

Kath Schamburg
Hendra Secondary College
Widdop Street
CLAYFIELD Q 4011

A. Prof Glenn Rowley
Monash University
Wellington Road
CLAYTON VIC 3168

Appendix 4 Means for each teacher in each school on each Key Competency

School 1

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	3.69	3.85	3.54	2.79	4.08	2.57	4.21	4.15	4.54
St Dev	1.18	1.21	1.33	0.8	1.19	1.09	1.42	1.34	1.56
Count	13	13	13	14	13	14	14	13	13
2 Average	3.6	4.92	3.45		4.4	4.7	3.4	4.08	5.08
St Dev	0.84	1.51	0.82		0.97	1.49	0.89	1.16	1.93
Count	10	12	11	0	10	10	5	12	12
3 Average	3.38	4	3.31	3.46	3	4.31	3.85	3.85	3.58
St Dev	1.76	2	1.93	1.85	1.87	1.7	1.63	1.95	2.23
Count	13	13	13	13	13	13	13	13	12
4 Average	4.45	4.95	4.14	3.77	5.05	4.86	4.23	4.18	4.55
St Dev	1.18	1.29	1.13	0.97	1.33	1.25	1.07	0.91	1.5
Count	22	22	22	22	22	22	22	22	22
5 Average	4.67	5.67	5.67	5	6	6	5	5	5
St Dev	2.08	0.58	1.53	2.83	1	1.73	1.41	2	1
Count	3	3	3	2	3	3	2	3	3
6 Average	4.8	4.8	4.4	4.1	5.4	5	5.1	4.9	5.3
St Dev	1.32	1.32	1.26	0.99	1.26	1.49	1.1	1.29	1.42
Count	10	10	10	10	10	10	10	10	10
7 Average	3.04	5.07	4		3.63	3.89	2	3.41	4.19
St Dev	1.09	1.27	1.57		1.33	1.69		1.37	1.36
Count	27	27	27	0	27	27	1	27	27
8 Average	4.57	4.71	4.18		4	3.43	5	3.57	4.54
St Dev	1.83	1.41	1.47		1.33	1.17		1.03	1.35
Count	28	28	28	0	28	28	1	28	28
9 Average	5	5.67	5.25	4.25	3.58	4.5	5.17	4.83	4.75
St Dev	1.54	1.23	1.48	0.97	1.83	1.73	1.47	1.75	1.71
Count	12	12	12	12	12	12	12	12	12
10 Average	5	5.2	5.2	4.4	3	4.3	4.1	3.9	4.89
St Dev	0.94	1.23	1.14	1.26	0.94	0.95	0.74	1.37	1.17
Count	10	10	10	10	10	10	10	10	9
11 Average	4.84	5.24	4.8		5.04	5	4.68	4.88	5.04
St Dev	1.18	0.97	1.38		0.79	1.76	0.99	1.45	1.51
Count	25	25	25	0	25	25	25	25	25

School 2

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	5.63	6.08	5.58	5.47	5.88	5.83	5.71	5.96	6.33
St Dev	1.01	1.02	1.1	0.87	0.95	1.01	0.69	0.81	0.76
Count	24	24	24	17	24	24	24	24	24
2 Average	4.05	4.42	3.89	3.89	4.84	3.63	4.37	3.95	4.47
St Dev	1.03	0.77	1.37	1.1	0.6	1.42	0.83	1.18	1.26
Count	19	19	19	19	19	19	19	19	19
3 Average	4.61	5	4.5	4.72	4.7	4.39	4.5	4.22	4.67
St Dev	1.14	1.33	1.38	1.32	1.06	1.38	1.04	0.94	1.28
Count	18	18	18	18	10	18	18	18	18
4 Average	4.78	4.69	4.67	4.7	5.21	4.56	4.44	4.67	4.59
St Dev	1.72	1.16	1.39	1.66	1.05	2.04	1.25	1.66	1.31
Count	27	26	27	27	14	27	27	27	27
5 Average	5.26	5.03	5.06		5.37	4.63		4.94	4.82
St Dev	1.24	1.22	1.28		1.35	1.59		1.54	1.29
Count	35	34	34	0	35	35	0	34	34
6 Average	4.65	5.35	4.95	4	4.95	5	4.89	4.95	4.7
St Dev	1.57	1.5	1.54	1.83	1.35	1.72	1.88	1.9	1.66
Count	20	20	20	4	19	20	19	20	20
7 Average	4.93	5.17	5.23		4.83	4.93	4.58	5.03	5.07
St Dev	1.41	1.09	1.33		1.34	1.44	1.08	1.5	1.41
Count	30	30	30	0	30	30	12	30	30
8 Average	5.12	5.54	5.12		5.38	4.96	5.18	5	5.65
St Dev	0.86	1.1	1.03		0.85	1.06	0.6	1.2	1.16
Count	26	26	26	0	26	25	11	26	26
9 Average	3.77	4.15	4.12	4.12		3.62	3.69	3.96	4.04
St Dev	1.18	0.67	1.03	0.65		0.9	0.62	1.04	1.04
Count	26	26	26	26	0	26	26	26	26
10 Average	4.47	4.61	4.63		4.36	4.75	7	4.48	4.32
St Dev	1.35	1.11	1.29		1.21	1.49		1.39	1.39
Count	57	57	57	0	56	55	1	56	56

School 3

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	5.5	6.13	5.13	5.25	5.29	6.13	5	6.13	6
St Dev	0.93	0.83	1.64	0.5	0.76	1.89	1.1	1.55	1.77
Count	8	8	8	4	7	8	6	8	8
2 Average	4	4.75	4.88	4.5	4.75	4.75	3.88	4.25	5.13
St Dev	1.69	1.39	1.46	1.77	2.05	2.19	1.64	1.83	1.81
Count	8	8	8	8	8	8	8	8	8
3 Average	5.46	5.38	5.69		6.08	5		4.92	5.31
St Dev	1.51	1.5	1.93		1.26	1.87		1.71	1.6
Count	13	13	13	0	13	13	0	13	13
4 Average	5.5	6		5.63	6.38	5.38		5.38	5.75
St Dev	1.41	1.2	1.77		1.19	1.41		1.77	1.04
Count	8	8	8	0	8	8	0	8	8
5 Average	4.83	5.83	5.33	4.92	5.67	4.83	5	5.08	6.17
St Dev	1.19	0.94	1.44	1.16	0.89	1.75	0	1	0.83
Count	12	12	12	12	12	12	12	12	12
6 Average	6	4.5	5	4.63	4.71	5.63	5.13	4.67	5.8
St Dev	1.07	0.93	1.41	1.3	1.38	1.19	1.55	0.82	1.1
Count	8	8	8	8	7	8	8	6	5
7 Average	5	5	4		5	2.5	2.5		3
St Dev	2.83	2.83	4.24		0	2.12		2.12	2.83
Count	2	2	2	0	2	2	0	2	2
8 Average	4.5	4.67	4.83	5.5	4.67	4.83	5.25	4.42	5.42
St Dev	1.57	1.56	1.4	1	1.61	1.7	1.16	1.88	1.51
Count	12	12	12	4	12	12	8	12	12
9 Average	5.42	5.42	5.5	4	4.27	4.36	2.82	4.5	5.27
St Dev	1.68	1.16	2.07	1.28	1.49	1.57	0.98	1.65	1.27
Count	12	12	12	12	11	11	11	10	11
10 Average	6	5.33	6.17	4.83	6.5	6.17	6	6.17	7
St Dev	2.28	1.97	1.94	1.94	1.97	2.32	2.37	2.32	2
Count	6	6	6	6	6	6	6	6	6
11 Average	5.5		6.4	5.5	6.1	5.8	5.8	5.7	5.9
St Dev	1.27		0.97	1.27	1.2	1.32	0.92	1.25	1.37
Count	10	10	10	0	10	10	10	10	10
12 Average	5.22	5.94	5.67	5.71	5.41	5.39	4.83	5.44	5.56
St Dev	1.83	1.11	1.28	1.4	1	1.2	1.38	1.46	1.25
Count	18	18	18	17	17	18	18	18	18

School 4

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	3.06	4.25	4.13	3.75	2.88	4.31	4.56	3.5	4.94
St Dev	1.73	1.57	1.75	1.53	0.96	1.92	1.21	1.51	1.69
Count	16	16	16	16	16	16	16	16	16
2 Average	3.64	4.52	4.54	3.48	4.52	4.2	4.4	4.44	4.04
St Dev	1.68	1.48	1.44	1.66	1.56	1.32	1.76	1.45	1.24
Count	25	25	24	25	25	25	25	25	25
3 Average	3.14	3.36	3.29	3.21	3.29	3.14	3.71	2.93	3.43
St Dev	2.21	1.45	1.59	1.97	0.73	1.99	1.54	1.94	1.34
Count	14	14	14	14	14	14	14	14	14
4 Average	4.25	3.83	5		4.5	4.42	4.17	4.08	4.17
St Dev	2.63	2.55	2.24		2.47	2.81	2.48	2.71	2.86
Count	12	12	11	0	12	12	12	12	12
5 Average	1.57	2.71	3	2	1.67	1.57	3.29	1.33	3.29
St Dev	0.98	0.76	1.1	1.41	1.03	0.98	0.76	0.82	1.38
Count	7	7	6	2	6	7	7	6	7
6 Average	5.85	6.15	5.38		6.54	6.23		6.23	6.69
St Dev	1.14	1.07	1.12		1.45	1.42		1.17	1.03
Count	13	13	13	0	13	13	0	13	13
7 Average	4.18	4.65	3.53	3	3.53	4	3.75	3.88	4.88
St Dev	1.59	1.11	1.62	1.15	1.28	1.7	1.61	1.54	1.36
Count	17	17	17	7	17	17	16	17	16
8 Average	2.75	4	3		3.29	3.13	3.75	2.14	3.25
St Dev	0.71	1.07	1.07		0.76	0.35	1.04	1.07	0.71
Count	8	8	8	0	7	8	8	7	8
9 Average	5.29	6.08	5.58	4.67	5.85	5.67	4.73	5.5	5.4
St Dev	1.44	1.38	1.56	1.5	1.21	1.37	1.62	1.51	1.58
Count	14	13	12	12	13	12	11	12	10
10 Average	4.63	5.5	4.75		5.5	4.94		4.88	4.88
St Dev	1.36	1.51	1.61		1.15	1.69		1.86	1.59
Count	16	16	16	0	16	16	0	16	16
11 Average	3.79	3.17	3.95		3.42	3.63	3.47	3.79	4.58
St Dev	1.44	1.47	1.72		1.22	1.64	1.47	1.47	1.8
Count	19	18	19	0	19	19	19	19	19

School 5

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	5	6.14	6	5.86	5.43	4.86	5.75	5.86	5.86
St Dev	1.29	0.69	1.29	0.9	0.98	0.69	0.96	1.07	1.77
Count	7	7	7	7	7	7	4	7	7
2 Average	4.8	4.9	4.9	5.4		4.9	4	3.9	5
St Dev	1.14	0.99	0.57	0.97		0.99	0.82	0.57	0.67
Count	10	10	10	10	0	10	10	10	10
3 Average	5.5	5.88	5.75	5.5	6.38	5.5	5.5	4.88	5.75
St Dev	1.31	1.46	1.16	1.41	0.74	1.85	1.07	1.46	1.39
Count	8	8	8	8	8	8	8	8	8
4 Average	4	6.5	5	3		6.5	5	6	6.5
St Dev	1.41	2.12	2.83	1.41		0.71	2.83	1.41	2.12
Count	2	2	2	2	0	2	2	2	2
5 Average	4.33	4.56	4.56	4.67	5.67	4.33	4.33	5.11	4.56
St Dev	2.24	1.33	1.94	1.32	1.32	2.69	2.5	1.45	1.67
Count	9	9	9	9	9	9	9	9	9
6 Average	5.2	4.36	4.71		3.64			4.47	4
St Dev	1.82	1.5	1.86		0.93			1.51	1.73
Count	15	14	14	0	14	0	0	15	15
7 Average	4.6	5.36	4.86	4.69	4.38	4.54	3.92	4.85	5.46
St Dev	2.16	1.45	1.99	1.38	1.71	1.85	1.32	1.52	1.66
Count	15	14	14	13	13	13	13	13	13
8 Average	4.73	4.25	4.25		4.88	4.63		4.19	4
St Dev	2.25	1.44	1.81		1.82	2.06		2.56	1.63
Count	15	16	16	0	16	16	0	16	16
9 Average	4.83	4.83	4.92	5.25	5.17	5.33	4.9	4.83	4.83
St Dev	1.11	1.34	1.31	1.82	0.94	1.23	1.1	0.94	1.03
Count	12	12	12	12	12	12	10	12	12
10 Average	6	5.08	5.58	5.64	5.85	5.85	7	6	6.92
St Dev	1.91	1.88	2.07	1.86	1.14	1.91	1.41	1.73	0.79
Count	13	12	12	11	13	13	4	13	12
11 Average	5.22	4.78	5		5.22	5		5.22	4.56
St Dev	1.56	1.56	2.24		1.56	2.24		1.56	1.33
Count	9	9	9	0	9	9	0	9	9
12 Average	4.5	3	4.5	5	4	3.5	5	5	4
St Dev	2.12	0	2.12		1.41	2.12			1.41
Count	2	2	2	1	2	2	1	1	2

School 6

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	5.26	5.68	5.89	4		6	5.26	5.37	
St Dev	1.19	1.2	1.6	1.41		1	0.93	0.83	
Count	19	19	18	2	0	19	19	19	0
2 Average	5.19	5.56	5.28	4.62	5.64	5.03	5.25	5.29	4.94
St Dev	1.39	1.25	1.54	1.71	1.2	1.34	1.26	1.41	1.64
Count	36	36	36	13	36	36	4	35	35
3 Average	5.51	5.44	5.74	5.43	5.28	5.1	5.28	5.56	
St Dev	1.64	1.64	1.63	1.64	1.72	1.26	1.83	1.65	
Count	43	43	43	0	42	43	41	43	43
4 Average	4.37	3.86	3.82	4.78	3.74	4.16	4.16	4.18	4.55
St Dev	1.81	1.55	1.5	1.41	2.02	1.67	1.97	1.56	1.59
Count	38	37	38	23	38	38	25	38	38
5 Average	5.78	4.59	4.84	3.95	4.76	5.07	4.14	5.28	5.43
St Dev	1.44	2.02	1.72	1.48	1.74	1.62	1.29	1.69	1.63
Count	46	46	43	41	46	46	14	43	46
6 Average	4.58	5.5	4.81	5	4.38	5.83	5.45	5.04	5.05
St Dev	1.14	1.15	1.33	0	1.2	1.59	1.37	1.78	1.81
Count	24	20	21	4	21	23	22	24	22
7 Average	5.06	5.78	5.89	5.11	6	4.78	5.75	5.12	5.56
St Dev	1.73	1	1.13	1.68	1	1.63	1.5	1.45	0.78
Count	18	18	18	18	3	18	4	17	18
8 Average	4.76	5	4.82	5		4.47	4.94	4.94	5.18
St Dev	1.15	0.94	1.24	1.22		1.59	1.03	1.2	1.13
Count	17	17	17	17	0	17	17	17	17

School 7

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	4.14	4.65	2.88	3.43	4.42	3.5	3.13	3.75	4.46
St Dev	2.01	1.61	2.09	2.04	1.89	2.3	1.68	2.09	2.02
Count	22	23	24	23	24	24	24	24	24
2 Average	3.83	5	3.67	3.5	4.42	4	3.08	3.67	4.08
St Dev	1.53	1.54	1.44	1.51	1.24	1.76	1.51	1.37	1.62
Count	12	12	12	12	12	12	12	12	12
3 Average	4.9	4.67	4.62	4.35	5.05	4.95	5.05	4.86	4.76
St Dev	1.51	1.35	1.32	1.17	1.1	1.6	1.28	1.65	1.79
Count	21	21	21	17	20	21	21	21	21
4 Average	6	5.76	5.71		5.59	5.82	5.5	5.63	5.71
St Dev	1.06	0.9	1.4		1.46	1.24	1	1.26	1.4
Count	17	17	17	0	17	17	4	16	17
5 Average	4	4	4.25	3	3	4	4.5	4.25	4
St Dev	2.24	2.24	2.5	0	0	2.24	1.91	2.5	2.24
Count	5	5	4	3	3	5	4	4	5
6 Average	5.25	4.75	5.3	5.7		5.5	5.45	5.35	6.8
St Dev	1.45	1.16	1.53	1.81		1.73	1.73	1.42	1.15
Count	20	20	20	20	0	20	20	20	20
7 Average	4.86	5.38	4.86	4	4.69	5.21	4.38	4.58	4.88
St Dev	1.48	1.59	1.48		1.31	1.4	1.65	1.3	1.8
Count	29	29	28	1	29	29	26	26	26
8 Average	5	5.21	5.21	5.21		5.57		5.79	6
St Dev	1.71	1.42	1.31	1.53		1.79		1.53	1.04
Count	14	14	14	14	0	14	0	14	14
9 Average	3.69	3.79	3.6	5	3.31	4	3	3.63	3.94
St Dev	0.95	0.89	0.83	0	0.79	1.32	0	0.96	1.34
Count	16	14	15	2	16	16	5	16	16

School 8

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	4.75	4.75	5	4.5	4.5	4.5	5.25	4	5.75
St Dev	0.5	0.96	0.82	0.58	1.29	1	0.5	0	0.96
Count	4	4	4	4	4	4	4	4	4
2 Average	5.08	5.75	5.33	5.5	4.75	5.17	4.83	5	5.5
St Dev	1.93	1.76	1.87	1.93	1.71	1.8	1.99	2.09	1.73
Count	12	12	12	12	12	12	12	12	12
3 Average	4.25	4.33	3.44		4.67	4.22	3	4.56	3.44
St Dev	2.05	1.94	2.3		1.73	1.99	1.73	2.19	1.94
Count	8	9	9	0	9	9	9	9	9
4 Average	4.33	4.87	4.53	5.75	4.67	4.53	6.14	4.47	4.36
St Dev	2.09	1.41	2.1	0.96	1.76	2.45	1.46	2.56	2.47
Count	15	15	15	4	15	15	7	15	14
5 Average	4.36	4.5	4.68	4.14	4.45	4.59	3.55	4.32	4.41
St Dev	1.43	1.22	0.89	1.52	1.06	1.5	1.01	1.39	1.3
Count	22	22	22	22	22	22	22	22	22
6 Average	4.73	5.09	4.73		4.91	4.64	4.7	4.73	5.27
St Dev	1.49	0.94	1.42		1.45	1.57	1.06	1.42	1.19
Count	11	11	11	0	11	11	10	11	11
7 Average	4.5	4.71	4.43	4.62	4.57	4.57	3.64	4.43	4.36
St Dev	1.29	1.44	1.34	1.26	1.22	1.4	0.93	1.45	1.22
Count	14	14	14	13	14	14	14	14	14
8 Average	5.19	5.5	5	4.14	5.13	5.19	4.75	4.88	5.19
St Dev	1.64	1.46	1.79	1.57	1.75	2.01	1.39	1.93	1.97
Count	16	16	16	7	16	16	16	16	16
9 Average	3.36	4.36	3.18		4	3.73	3	3.73	3.9
St Dev	2.06	1.91	1.99		2.19	2.2	1.79	1.68	1.52
Count	11	11	11	0	11	11	11	11	10
10 Average	3.9	4.4	3.8	3.4	4	3.8	4.4	3	5
St Dev	2.42	1.51	2.53	1.51	1.41	1.81	1.84	1.25	1.63
Count	10	10	10	10	10	10	10	10	10
11 Average	5.89	6.16	6			5.8	6	5.79	6.48
St Dev	1.23	1.07	1.05			1.11	1.15	1.23	1.03
Count	18	19	19	0	0	20	19	19	21
12 Average	4.75	5.67	4.67	3.27	5.08	5.25	3.14	5.17	6
St Dev	1.54	1.23	1.37	1.49	1.24	1.14	1.46	1.27	1.13
Count	12	12	12	11	12	12	7	12	12

School 9

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	4.5	5.5	4	4.38	4.38	4	4.88	4.43	4
St Dev	1.41	0.93	1.51	1.19	1.19	1.85	1.36	1.9	1.07
Count	8	8	8	8	8	8	8	7	8
2 Average	3.86	4.14		4.43		3.29		2.29	3
St Dev	2.04	1.46		1.4		1.89		1.5	1.53
Count	7	7	0	7	0	7	0	7	7
3 Average	3.58	3.32	3.89			3.84		3.74	3.26
St Dev	1.57	1.34	1.66			1.86		1.85	2.05
Count	19	19	19	0	0	19	0	19	19
4 Average	4.38	4.48	4.05	4.89	4.4	3.95	4.29	4.29	4.86
St Dev	1.72	1.33	1.5	1.9	0.89	1.56	1.65	1.23	1.49
Count	21	21	21	9	5	21	21	21	21
5 Average	3.62	3.85	3.54	3.58		3.69	4.15	3.85	4.38
St Dev	0.96	0.55	0.88	0.67		1.11	0.99	0.69	1.04
Count	13	13	13	12	0	13	13	13	13
6 Average	5.15	4	4.87		4.73	4.6		4.53	4.2
St Dev	1.07	1.6	1.25		1.28	1.35		1.36	1.32
Count	13	12	15	0	15	15	0	15	15
7 Average	5.36	6	5.43	5.86	4.43	4.86	5.21	4.29	5.43
St Dev	1.6	1.04	1.28	1.35	1.02	1.7	0.97	1.07	1.16
Count	14	14	14	14	14	14	14	14	14
8 Average	5.31	5.38	4.77		5.15	5	4.54	5.08	5.38
St Dev	1.32	1.39	1.64		1.41	1.29	1.2	1.55	1.61
Count	13	13	13	0	13	13	13	13	13
9 Average	4.13	4.71	4.14	5		4.43	4.29	4.43	4.29
St Dev	1.55	0.95	1.35	1.91		1.51	1.25	1.51	0.76
Count	8	7	7	7	0	7	7	7	7
10 Average	4.5	4.63	4.69	4.63	4.14	4.47	4.6	4.31	3.69
St Dev	1.55	1.09	1.35	1.09	1.03	1.41	1.12	1.49	1.3
Count	16	16	16	16	14	15	15	16	16

School 10

Teacher	KC1	KC2a	KC2b	KC3	KC4	KC5	KC6	KC7	KC8
1 Average	5.74	6.05	5.58		4.95	5.84	5.53	6.37	6.26
St Dev	1.15	0.85	1.07		1.27	1.17	0.84	1.3	0.87
Count	19	19	19	0	19	19	19	19	19
2 Average	5.29	4.97	5.03	6	5.12	5.29	5.41	5.36	5.64
St Dev	1.27	1.51	1.24		1.05	1.03	0.78	1.19	0.96
Count	34	33	33	1	33	34	34	33	33
3 Average	4.26	5.09	4.74	4.43	4.78	4.87	5	5.27	5.36
St Dev	1.79	1.38	1.71	1.16	1.59	1.52	0.93	1.2	1.53
Count	23	23	23	23	23	23	22	22	22
4 Average	5.44	5.94	5.31		5.39	5.28		5.25	4.8
St Dev	1.11	1.13	0.9		1.05	1.4		1.14	1.19
Count	32	32	32	0	31	32	0	32	30
5 Average	5.42	6.08	5.75	5.29	5.83	6.08	5.92	5.96	6
St Dev	1.38	1.32	1.26	1.04	1.34	1.41	1.41	1.52	1.14
Count	24	24	24	24	24	24	24	24	24
6 Average	5.11	5.67	5.28	5.39		5.67	5.89	6	6.28
St Dev	1.71	1.78	1.74	1.69		1.81	1.23	1.68	1.45
Count	18	18	18	18	0	18	18	18	18
7 Average	5.43	5.76	5.52	5.52	5.62	5.62	5.57	5.48	5.81
St Dev	1.29	1.26	1.36	1.21	1.02	1.4	1.08	1.33	1.33
Count	21	21	21	21	21	21	21	21	21
8 Average	4.5	4.95	4.32	3.82	5.05	4.73	5.09	5	5.27
St Dev	0.86	0.72	0.95	1.01	0.58	0.83	0.43	0.82	0.77
Count	22	22	22	22	22	22	22	22	22
9 Average	5.5	5.4	4.88	5.62	5.56	5.65	5.77	5.69	5.96
St Dev	1.36	1.22	1.66	1.44	1.16	1.26	0.95	1.32	1.04
Count	26	25	26	26	25	26	26	26	26
10 Average	4.79	6.21	4.71	5.14	5.43	5.93		5.57	6.14
St Dev	1.37	0.97	1.49	1.46	1.4	1.21		1.5	1.23
Count	14	14	14	14	14	14	0	14	14
11 Average	5.96	5.68	5.76	6.04	5.56	6.32	5.92	6.32	5.96
St Dev	1.02	0.99	1.05	1.02	0.87	0.99	0.86	0.99	0.89
Count	25	25	25	25	25	25	25	25	25

Appendix 5 An analysis of teacher leniency or harshness and consistency

School 1

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	10	-0.53	24	0.5	10	-0.32	24	0.37	10	-0.83	24	0.64
2	22	-0.21	41	0.36	22	-0.33	42	0.22	22	0.16	41	0.33
3	12	-0.54	25	0.55	12	-0.69	25	0.38	12	-0.67	25	0.77
4	25	-0.51	43	0.37	25	-0.03	44	0.19	25	-0.5	44	0.32
5	3	-1.53	6	1.15	3	-1.46	6	0.58	3	-1.75	6	0.9
6	13	1.40	24	0.68	13	1.22	24	0.32	13	1.63	24	0.54
7	27	0.79	25	0.28	27	0.38	25	0.66	27	-0.39	25	0.71
8	28	-0.02	58	0.37	28	0.50	58	0.24	28	0.45	58	0.37
9	10	-0.79	17	0.17	10	-0.55	18	0.21	10	-0.23	17	0.43
10	12	1.01	25	0.47	12	0.14	29	0.54	12	1.23	26	0.68
11	15	0.94	32	0.29	15	1.13	33	0.29	15	0.89	32	0.34
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	10	-0.15	11	0.84	10	1.55	24	0.56	10	0.16	24	0.85
2	22	-0.31	19	0.36	22	-0.08	41	0.33	22	-0.7	41	0.54
3	12	-0.67	9	1.06	12	0.55	25	0.5	12	-0.02	25	0.86
4	24	0	0	0	25	-0.58	43	0.38	25	-0.73	43	0.47
5	3	-0.19	3	1.3	3	-2.57	6	0.37	3	-1.23	6	0.54
6	13	0.77	13	0.41	13	1.97	24	0.49	13	0.49	24	0.55
7	27	0	0	0	27	0.03	25	0.7	27	-0.2	25	0.67
8	23	0	0	0	28	0.6	58	0.31	28	1.36	59	0.41
9	10	-0.7	13	0.07	10	-1.03	17	0.32	10	-0.9	17	0.57
10	12	0	0	0	12	0.09	25	0.32	12	-0.16	25	0.4
11	15	1.25	18	0.19	15	0.18	32	0.45	15	1.93	33	0.39
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	10	0.67	16	0.72	10	0.56	24	0.64	10	0.12	22	0.65
2	22	-0.21	27	0.58	22	-0.2	42	0.46	22	0.04	42	0.53
3	12	-0.52	16	0.38	12	-0.38	25	0.51	12	0.05	25	0.61
4	25	0	20	0.53	25	-0.61	44	0.27	25	-0.4	42	0.44
5	3	-2.01	2	0.49	3	-1.07	6	0.33	3	-0.8	6	0.18
6	13	0.74	18	0.52	13	0.9	24	0.36	13	1.33	22	0.48
7	27	0	0	0	27	0.66	25	0.36	27	0.29	25	0.51
8	16	0.33	2	0.34	28	1.03	58	0.33	28	0.54	56	0.31
9	10	-1.16	15	0.58	10	-1.2	18	0.22	10	-1.1	18	0.81
10	12	1.66	11	0.49	12	0.3	29	0.46	12	-0.31	29	0.67
11	15	0.51	21	0.66	15	0.03	33	0.24	15	0.24	33	0.37

School 2

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	27	0.22	78	0.2	27	0.29	73	0.27	27	0.27	78	0.24
2	18	-0.07	63	0.2	18	0.03	63	0.36	18	0.07	63	0.31
3	26	-0.11	81	0.18	26	-0.37	81	0.21	26	-0.14	81	0.19
4	35	-0.75	112	0.16	35	-0.13	110	0.12	35	-0.4	112	0.13
5	26	0.56	84	0.21	26	0.57	83	0.21	26	0.35	84	0.27
6	20	-0.21	45	0.18	20	-0.52	45	0.3	20	-0.43	45	0.35
7	24	-0.85	83	0.15	24	-1.19	82	0.16	24	-0.68	83	0.16
8	30	-0.28	97	0.12	30	-0.28	96	0.13	30	-0.54	97	0.17
9	57	0.25	182	0.09	57	0.41	180	0.12	57	0.09	182	0.11
10	19	1.24	53	0.18	19	1.19	53	0.23	19	1.41	53	0.27
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	27	0.08	25	0.12	27	0.14	31	0.32	27	0.29	75	0.26
2	18	-0.42	17	0.24	10	0.37	31	0.4	18	0.07	61	0.33
3	22	0	0	0	26	-0.1	68	0.24	26	0.02	75	0.23
4	35	0	0	0	35	-0.46	87	0.15	35	-0.27	111	0.2
5	36	0.11	15	0.31	20	-0.29	33	0.19	26	0.69	84	0.32
6	20	-0.01	3	0.75	24	-0.87	62	0.17	20	-0.58	45	0.34
7	17	-1.02	16	0.19	30	0	71	0.22	24	-1.26	82	0.19
8	19	1.27	18	0.33	56	0.59	139	0.12	30	-0.46	96	0.13
9					19	0.62	50	0.24	56	-0.09	177	0.07
10									19	1.6	52	0.29
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	27	0.71	40	0.16	27	0.26	77	0.18	27	0.32	77	0.23
2	18	0.48	28	0.41	18	0.27	61	0.28	18	0.11	61	0.35
3	24	0.14	18	0.34	26	-0.06	80	0.13	26	-0.44	80	0.28
4	35	0	0	0	35	-0.46	109	0.19	35	-0.08	109	0.17
5	26	1.07	36	0.16	26	0.42	84	0.26	26	0.59	84	0.25
6	20	-0.06	21	0.31	20	-0.54	45	0.39	20	-0.32	45	0.45
7	24	-0.76	37	0.35	24	-1.18	83	0.15	24	-1.43	83	0.18
8	12	-0.29	19	0.41	30	-0.4	97	0.06	30	-0.27	97	0.13
9	1	-2.22	3	0.57	56	0.25	178	0.05	56	0.53	178	0.13
10	19	0.93	28	0.39	19	1.44	52	0.26	19	0.99	52	0.41

School 3

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	12	-0.63	31	0.29	12	-0.43	31	0.14	12	-0.35	31	0.42
2	12	1.25	30	0.44	12	0.17	30	0.22	12	0.81	30	0.4
3	8	1.29	16	0.76	8	0.94	16	0.52	8	0.21	16	0.44
4	8	-0.01	14	0.42	8	-0.47	14	0.51	8	-0.52	14	0.65
5	6	0.03	15	0.49	6	0.34	15	0.4	6	0.02	15	0.26
6	10	-0.55	25	0.25	10	-1.13	25	0.54	10	-0.1	25	0.41
7	8	-1.45	17	0.73	8	-1.73	17	0.26	8	-1.25	17	0.51
8	13	0.06	31	0.42	13	-0.11	31	0.43	13	0.01	31	0.3
9	12	0.15	29	0.44	12	0.75	29	0.77	12	-0.1	29	0.3
10	8	-0.76	18	0.46	8	0.31	18	0.43	8	0.45	18	0.51
11	18	0.24	44	0.45	18	-0.42	44	0.35	18	-0.19	44	0.42
12	2	0.38	6	0.5	2	1.79	6	1.42	2	0.99	6	0.46
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	12	0.88	15	0.4	12	0.94	27	0.53	12	0.08	28	0.35
2	12	0.86	18	0.43	12	0.37	30	0.57	12	1.21	30	0.7
3	8	-0.26	10	0.27	8	0.26	16	1.13	8	0.1	16	0.71
4	6	0.6	9	0.27	8	-0.94	13	0.48	8	0.16	15	0.91
5	4	-1.12	5	0.8	6	-0.84	15	0.74	6	-0.24	15	0.62
6	6	-0.91	4	0.68	10	-0.76	24	0.63	10	-0.88	25	0.36
7	8	0.4	7	0.55	7	-1.03	14	0.68	8	-2.32	17	0.62
8	18	-0.46	24	0.43	13	-0.47	31	0.48	13	0.33	30	0.5
9					12	0.51	29	0.84	12	-0.21	29	0.69
10					8	0.7	15	0.44	8	-0.62	18	0.48
11					18	-0.01	40	0.47	18	0.05	43	0.54
12					2	1.26	6	1.01	2	2.33	6	0.77
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	12	1.86	21	0.19	12	0.25	25	0.45	12	-0.31	27	0.54
2	12	0.446	22	0.44	12	0.71	30	0.51	12	-0.04	30	0.53
3	8	0.57	11	0.5	8	0.76	16	0.84	8	0.72	16	0.94
4	1	0	0	0	8	0.14	13	0.82	8	0.33	13	0.78
5	6	-1.2	11	0.9	6	-0.77	15	0.59	6	-0.51	15	0.75
6	10	-1.25	17	0.35	10	-0.87	23	0.33	10	-0.83	24	0.37
7	7	-0.35	11	0.43	8	-2.21	15	0.73	8	-1.4	16	1.05
8	9	-0.1	16	0.37	13	0.14	30	0.47	13	0.11	29	0.41
9	8	-0.69	10	0.38	12	-0.27	27	0.55	12	-0.24	27	0.87
10	18	0.69	31	0.27	8	0.55	12	0.79	8	-0.73	10	0.54
11					18	-0.4	40	0.34	18	0.1	41	0.44
12					2	1.97	6	0.67	2	2.79	6	0.99

School 4

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	19	0.72	32	0.4	19	1.41	30	0.29	19	0.6	29	0.44
2	14	-0.07	28	0.39	14	-0.87	26	0.38	14	-0.42	23	0.39
3	25	1.2	41	0.27	25	0.52	40	0.31	24	0.55	36	0.51
4	14	0.26	19	0.88	14	0.29	19	0.62	14	0.82	18	1.11
5	13	-2.78	24	0.45	13	-2.04	24	0.6	13	-1.6	24	0.28
6	12	0.95	23	0.62	12	0.92	23	0.39	12	0.32	21	0.4
7	16	0.48	24	0.74	16	0.04	24	0.35	16	0.8	24	0.55
8	17	-0.02	26	0.55	17	0.14	25	0.57	17	0.5	24	0.54
9	7	1.05	10	0.59	7	0.82	10	0.71	6	-0.18	8	0.91
10	16	-1.41	22	0.49	16	-1.31	22	0.55	16	-0.87	21	0.48
11	8	-0.38	9	0.25	8	0.07	9	0.52	8	0.36	8	0.12
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	19	0	0	0	19	0.58	31	0.17	19	1.35	31	0.45
2	14	-0.39	9	0.09	14	-1.44	27	0.37	14	-0.37	24	0.71
3	25	0.69	19	0.39	25	-0.01	41	0.2	25	1.15	40	0.29
4	14	0.01	6	1.23	14	0.9	19	0.52	14	0.34	19	0.72
5	16	-0.65	10	0.66	13	-2.86	24	0.41	13	-2.85	24	0.84
6	16	0.35	2	2.34	12	-0.02	23	0.26	12	1.08	22	0.41
7	2	0	0	0	16	1.03	24	0.53	16	-0.11	24	0.54
8	12	0	0	0	17	0.73	25	0.36	17	0.39	25	0.69
9	8	0	0	0	6	2.12	8	0.54	7	0.83	10	0.93
10					16	-1.41	21	0.26	16	-1.43	22	0.775
11					8	0.38	7	0.49	8	-0.38	9	0.86
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	19	1.22	26	0.68	19	0.89	31	0.44	19	0.18	30	0.58
2	14	0.24	18	0.55	14	-0.33	24	0.42	14	-0.31	20	0.47
3	25	-0.17	34	0.57	25	0.43	40	0.32	25	1.26	39	0.37
4	14	-0.21	13	0.66	14	0.48	19	0.5	14	0.1	19	0.67
5	12	0.25	21	0.64	13	-2.91	24	0.61	13	-2.73	24	0.78
6	16	-0.67	19	0.38	12	0.91	22	0.42	12	1.38	20	0.29
7	17	0.02	17	0.66	16	0.01	24	0.48	16	-0.76	24	0.47
8	7	-0.52	6	0.8	17	0.3	25	0.38	17	0.1	25	0.33
9	12	0	0	0	6	1.53	8	1.05	7	0.81	10	1.18
10	8	-0.16	4	0.69	16	-1.55	20	0.87	16	-0.87	22	0.62
11					8	0.23	7	0.63	8	0.84	9	0.92

School 5

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	15	0.08	35	0.32	15	-0.43	31	0.35	15	0.02	32	0.29
2	15	-0.77	32	0.45	15	0.35	29	0.56	15	-0.21	30	0.54
3	12	0.96	24	0.33	12	0.91	24	0.74	12	0.89	24	0.71
4	7	-0.69	15	0.65	7	-1.32	15	0.7	7	-1.1	15	0.75
5	11	-1.10	14	0.87	11	-1.54	14	0.93	11	-1.27	14	0.85
6	2	2.40	4	0.32	2	-0.54	4	0.39	2	1.16	4	0.14
7	8	-1.66	5	1.46	8	0.88	6	0.35	8	-1.23	6	1.3
8	16	0.27	38	0.43	16	0.74	38	0.32	16	0.72	39	0.32
9	13	-0.61	23	0.35	13	0.05	21	0.37	13	-0.02	20	0.6
10	9	0.41	15	0.54	9	0.09	15	0.59	9	0.29	15	0.52
11	9	-0.27	17	0.36	9	0.52	16	0.54	9	0.01	16	0.88
12	10	0.99	14	0.76	10	0.3	15	0.4	10	0.75	15	0.93
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	15	-0.05	14	0.24	15	0.95	28	0.28	15	0.18	26	0.53
2	12	0.78	17	0.25	14	1.09	26	0.38	12	0.79	23	0.47
3	7	-1.58	9	0.19	12	0.95	20	0.66	7	-0.02	13	0.73
4	9	-0.73	9	0.3	7	-0.31	15	0.61	11	-0.29	12	0.48
5	2	1.95	2	0	11	-1.68	14	0.64	2	-1.32	2	0
6	6	-2.15	2	0.58	8	-0.59	5	0.33	8	-0.08	5	1.28
7	11	0.5	13	0.41	16	0.48	37	0.42	16	0.39	33	0.17
8	9	-0.17	7	0.29	13	-0.67	18	0.54	13	-0.3	19	0.43
9	10	1.45	9	0.4	9	-0.55	15	0.56	9	0.22	13	0.6
10					9	0.34	16	0.51	9	-0.69	13	0.71
11									10	1.12	13	0.73
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	15	0.62	11	0.22	15	0.4	31	0.24	15	-0.07	30	0.33
2	10	0.54	11	0.3	15	0.1	30	0.44	15	1	29	0.68
3	4	0.11	4	0.86	12	1.11	24	0.59	12	1.01	24	0.51
4	11	-0.78	6	0.58	7	-1.59	15	0.68	7	-0.6	15	0.47
5	2	-0.38	2	0	11	-0.4	14	0.9	11	-1.19	14	0.55
6	5	-0.84	2	0.95	2	0	4	0.6	2	-0.03	4	0.53
7	4	-1.18	4	0.24	8	-2.45	4	0.3	8	-1.23	6	1.04
8	9	0.84	5	0.69	16	1.11	38	0.59	16	1.08	38	0.25
9	10	1.06	7	0.26	13	-0.31	23	0.28	13	-1.5	19	0.47
10					9	-0.19	15	0.59	9	0.13	14	0.33
11					9	-0.11	16	0.65	9	1.17	16	0.34
12					10	2.34	14	0.42	10	0.23	15	0.91

School 6

	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	38	1.11	113	0.31	38	1.61	108	0.15	38	1.79	110	0.3
2	17	-0.24	50	0.55	17	-0.29	49	0.4	17	-0.32	45	0.45
3	18	0.91	52	0.28	18	0.4	52	0.32	18	0.32	51	0.28
4	36	0.09	103	0.27	36	-0.19	100	0.27	36	0.02	101	0.19
5	19	-0.54	55	0.39	19	-0.74	53	0.21	18	-1.05	51	0.35
6	24	0.02	66	0.3	24	-0.53	54	0.5	21	-0.21	57	0.36
7	46	-0.67	127	0.23	46	0.51	124	0.27	46	0.39	115	0.27
8	43	-0.66	120	0.26	43	-0.52	116	0.23	43	-0.94	114	0.3
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	23	0.36	37	0.32	38	1.27	81	0.29	38	1.25	112	0.29
2	17	-0.81	20	0.68	3	1.23	9	0.14	17	0.18	50	0.38
3	18	0.31	28	0.2	36	-0.88	80	0.23	18	1.21	52	0.39
4	13	0.01	23	0.22	21	-0.31	46	0.42	36	0.2	103	0.24
5	2	-0.72	2	0.73	46	-0.25	89	0.32	19	-1.03	54	0.21
6	4	0.32	4	0.96	43	-1.07	81	0.34	23	-1.29	63	0.17
7	41	0.54	46	0.25					46	-0.08	126	0.27
8									43	-0.43	120	0.12
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	25	0.77	43	0.51	38	1.27	110	0.37	38	1.06	99	0.24
2	17	-0.35	29	0.42	17	-0.52	50	0.44	17	-0.29	48	0.37
3	4	0	7	0.76	17	0.9	47	0.36	18	0.26	48	0.21
4	4	0.08	9	0.49	35	-0.06	99	0.36	35	0.29	92	0.13
5	19	-0.26	32	0.23	19	-0.33	55	0.47	22	-0.28	53	0.39
6	23	-1.12	36	0.44	24	-0.62	65	0.26	46	-0.43	113	0.14
7	14	1.18	25	0.64	46	-0.24	120	0.18	43	-0.59	107	0.15
8	41	-0.3	61	0.24	43	-0.42	118	0.25				

School 7

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	22	0.31	38	0.34	22	-0.31	35	0.36	22	1.4	39	0.34
2	20	-0.07	45	0.08	20	0.39	45	0.25	20	-0.47	46	0.17
3	14	0.69	31	0.46	14	0.49	31	0.35	14	0.39	32	0.3
4	29	-0.14	48	0.14	29	-0.5	47	0.18	28	-0.36	48	0.23
5	17	-1.20	29	0.35	17	-0.86	28	0.37	17	-0.69	28	0.45
6	21	-0.45	40	0.3	21	-0.10	40	0.24	21	-0.5	40	0.33
7	12	-0.50	10	0.66	12	-0.74	10	0.82	12	-0.13	10	0.37
8	16	1.16	32	0.61	16	1.35	29	0.35	16	0.98	31	0.54
9	5	0.21	5	1.17	5	0.28	5	1.23	5	-0.62	4	0.7
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	23	1.09	20	0.36	24	-0.01	24	0.33	24	0.89	40	0.33
2	20	-0.85	25	0.29	20	0	0	0	20	-0.06	46	0.31
3	14	0.8	18	0.31	14	0	0	0	14	0.49	32	0.31
4	29	-0.38	1	0	29	0.17	31	0.18	29	-0.55	48	0.33
5	21	0.46	15	0.29	17	-0.61	20	0.43	17	-1.01	29	0.42
6	12	-0.87	4	0.17	21	-0.79	26	0.2	21	-0.67	40	0.31
7	16	-0.87	2	0	12	-0.84	9	0.21	12	-0.86	10	0.52
8	5	0.62	1	0	16	1.39	24	0.28	16	0.93	32	0.39
9					5	0.69	2	0.52	5	0.84	5	1.06
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	24	1.66	27	0.38	24	0.52	40	0.16	24	0.11	40	0.34
2	20	-0.27	32	0.17	20	0	44	0.2	20	-1.3	44	0.22
3	14	0	0	0	14	0.09	32	0.25	14	-0.28	32	0.29
4	29	-0.04	32	0.27	29	0.09	42	0.33	29	0.61	42	0.35
5	5	-1.02	8	0.3	16	-0.84	27	0.46	17	-0.7	28	0.47
6	21	-0.54	32	0.29	21	-0.74	38	0.19	21	-0.24	38	0.29
7	12	-0.71	5	0.49	12	0.35	9	1.11	12	-0.18	9	0.51
8	15	1.18	6	0.63	16	1.14	32	0.32	16	1.25	32	0.63
9	5	-1.68	2	1.14	5	-0.61	4	0.78	5	0.73	5	1.8

School 8

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	22	-0.27	14	0.21	22	-0.81	15	0.23	22	-0.87	15	0.54
2	11	-0.27	22	0.68	11	0.05	23	0.46	11	-0.43	23	0.69
3	4	0.6	7	0.57	4	1.3	87	0.449	4	1.06	7	0.38
4	15	0.6	32	0.57	15	0.58	33	0.39	15	0.14	33	0.54
5	12	-0.05	21	0.66	12	-0.61	21	0.66	12	-0.88	21	0.68
6	11	0.21	19	0.35	11	-0.29	20	0.32	11	0.77	20	0.41
7	16	-0.56	33	0.17	16	-0.68	32	0.24	16	-0.5	32	0.32
8	12	-0.46	25	0.45	12	-1.06	25	0.42	12	-0.47	25	0.57
9	22	-0.02	41	0.59	22	0.50	41	0.22	22	-0.39	41	0.61
10	8	-0.04	15	0.73	9	0.00	16	0.51	9	1	16	1.06
11	14	0.36	28	0.55	14	0.51	28	0.32	14	0.22	28	0.45
12	10	0.31	15	0.82	10	0.61	15	0.53	10	0.35	15	0.9
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	4	0.65	6	0.28	11	-0.07	21	0.36	22	-0.73	17	0.3
2	4	0.19	8	0.83	4	1.35	7	0.11	11	0.09	23	0.65
3	12	-1.62	7	0.74	15	0.21	30	0.15	4	1.12	7	0.75
4	7	-0.09	6	0.58	12	0.32	20	1.04	15	0.58	33	0.92
5	11	0.68	9	0.33	11	-0.81	20	0.52	12	-0.5	21	1.09
6	22	-0.14	21	0.24	16	-0.61	31	0.45	11	-0.13	20	0.63
7	14	0.08	13	0.36	12	-0.76	21	0.4	16	-0.35	33	0.38
8	10	0.26	8	0.44	22	-0.06	39	0.34	12	-1.11	26	0.59
9					9	-0.31	16	0.65	22	-0.21	41	0.43
10					14	0.37	26	0.63	9	-0.06	16	0.63
11					10	0.37	15	0.21	14	-0.43	28	0.43
12									10	-0.76	15	0.89
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	22	-1.17	14	0.47	22	-0.69	17	0.47	22	-1.68	16	0.66
2	11	-0.58	19	0.69	11	-0.16	23	0.54	11	-0.14	23	0.51
3	4	-0.54	7	0.78	4	1.06	7	0.83	4	0.52	7	0.75
4	7	-1.29	18	1.02	15	0.55	35	1.14	15	1.09	30	0.56
5	12	-0.98	21	1.27	12	-0.65	21	1.44	12	0.15	21	0.84
6	11	0.89	18	0.74	11	-0.19	20	0.65	11	-0.48	18	0.58
7	16	-0.4	28	0.53	16	-0.2	33	0.55	16	-0.34	33	0.39
8	7	1.19	14	0.61	12	-0.85	26	0.54	12	-1.66	25	0.39
9	22	0.81	33	0.5	22	-0.07	41	0.55	22	0.37	40	0.54
10	9	1.17	15	0.96	9	-0.33	16	0.91	9	1.13	16	0.8
11	14	0.94	28	0.37	14	0.21	28	0.44	14	1.15	27	0.46
12	10	-0.04	13	1.03	10	1.32	15	0.72	10	-0.13	14	0.73

School 9

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	3	0	0	0	3	0	0	0	3	0	0	0
2	13	-0.91	34	0.44	13	-0.85	33	0.31	13	-0.28	32	0.26
3	21	0.46	49	0.32	21	0.59	50	0.25	21	0.89	46	0.15
4	8	-0.52	20	0.52	8	-1.39	19	0.56	8	-0.47	19	0.32
5	14	-0.51	25	0.45	14	-1.16	24	0.32	14	-0.57	24	0.48
6	19	0.71	47	0.27	19	1.23	45	0.39	19	0.46	42	0.17
7	8	0.15	17	0.46	8	0.12	16	0.65	8	0.21	15	0.19
8	16	-0.19	40	0.41	16	-0.01	40	0.25	16	-0.25	39	0.26
9	15	-0.59	27	0.24	15	0.57	24	0.77	15	-0.77	30	0.37
10	7	0.41	18	0.72	7	0.34	17	0.99	7	25	0	0
11	13	0.98	25	0.6	13	0.56	24	0.25	13	0.78	25	0.28
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	13	0	0	0	13	-0.32	14	0.29	3	0	0	0
2	21	-0.09	8	0.51	21	0.53	7	0.76	13	-0.58	34	0.3
3	8	-0.39	9	0.35	8	-0.07	9	0.72	21	0.7	50	0.13
4	14	-0.83	11	0.65	14	-0.2	11	0.41	8	-0.78	20	0.34
5	18	0	0	0	19	0	0	0	14	-0.26	25	0.41
6	8	-0.28	8	0.32	8	0	0	0	19	0.32	45	0.14
7	16	0.33	16	0.32	16	0.72	15	0.4	8	-0.12	16	0.48
8	7	0.25	7	0.62	15	-0.66	14	0.46	16	0.05	38	0.28
9	13	1.01	13	0.21	7	0	0	0	15	-0.69	31	0.49
10					13	0	0	0	7	0.69	16	0.6
11									13	0.67	25	0.31
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	3	0	0	0	3	0	0	0	3	0	0	0
2	13	0.12	23	0.36	13	-0.89	34	0.24	13	-0.87	34	0.37
3	21	0.28	30	0.41	21	0.33	49	0.17	21	-0.15	50	0.23
4	8	-0.83	10	0.71	8	-0.91	17	0.75	8	-0.67	20	0.72
5	14	-0.61	16	0.37	14	0.1	25	0.56	14	-0.69	25	0.59
6	19	0	0	0	19	0.18	46	0.45	19	1.06	46	0.53
7	8	0.17	11	0.63	8	-0.17	17	0.29	8	0.38	17	0.48
8	16	-0.12	26	0.32	16	0	40	0.41	16	0.68	41	0.28
9	7	0	0	0	15	-0.57	31	0.37	15	-0.63	31	0.35
10	13	0	16	0.48	7	1.39	17	0.38	7	1.24	17	0.45
11					13	0.54	24	0.58	13	-0.35	25	0.49

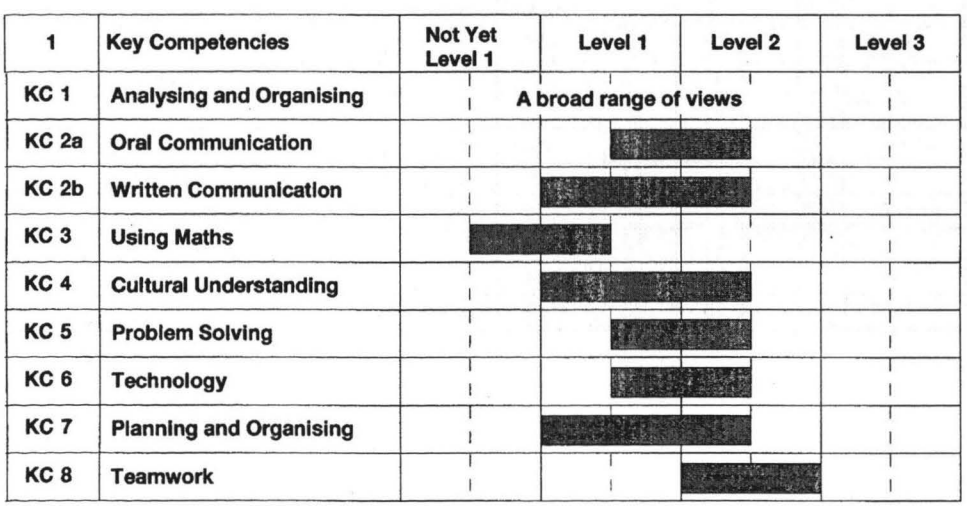
School 10

	KC1				KC2a				KC2b			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	21	-0.36	53	0.32	21	-0.29	53	0.17	21	-0.37	53	0.29
2	25	-0.08	84	0.17	25	0.41	83	0.12	25	-0.07	84	0.12
3	26	-0.36	87	0.09	26	0.71	84	0.1	26	0.79	87	0.14
4	18	-0.62	42	0.5	18	-0.61	41	0.28	18	-0.74	41	0.57
5	24	-0.32	58	0.22	24	-0.68	58	0.32	24	-0.53	58	0.2
6	23	0.74	65	0.31	23	0.41	63	0.24	23	-0.4	64	0.24
7	14	-0.54	43	0.22	14	-1.15	43	0.26	14	-0.36	43	0.31
8	34	-0.01	93	0.21	34	0.54	90	0.24	34	0.1	90	0.28
9	19	0.13	58	0.25	19	0.14	58	0.19	19	0.16	95	0.28
10	32	0.32	96	0.21	32	0.03	94	0.27	32	0.44	95	0.19
11	22	0.37	63	0.28	22	0.48	63	0.33	22	0.45	63	0.34
	KC3				KC4				KC5			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	21	-0.33	31	0.24	21	-0.35	52	0.38	21	-0.38	53	0.35
2	25	-0.11	52	0.25	25	0.07	83	0.17	25	-0.15	84	0.15
3	26	0.32	54	0.38	26	0.12	84	0.11	26	0.51	87	0.3
4	18	-0.45	20	0.37	18	0	0	0	18	-0.52	42	0.68
5	24	-0.03	37	0.3	24	-0.58	54	0.42	24	-0.73	58	0.29
6	23	0.88	47	0.38	23	0.32	58	0.32	23	0.69	65	0.43
7	14	-0.63	30	0.59	14	-0.8	40	0.43	14	-1.02	43	0.42
8	34	-0.89	2	0.21	34	0.18	77	0.32	34	0.31	93	0.32
9	19	0	0	0	19	0.68	52	0.22	19	0.12	58	0.34
10	32	0	0	0	32	0.19	88	0.22	32	0.66	96	0.25
11	22	1.25	47	0.32	22	0.16	56	0.23	22	0.52	63	0.37
	KC6				KC7				KC8			
	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.	Nos.	Adj.	Nos.	Const.
1	21	-0.25	39	0.43	21	0.01	52	0.3	21	-0.24	51	0.3
2	25	0	63	0.18	25	-0.09	84	0.15	25	0.17	83	0.21
3	42	0.14	65	0.19	26	0.54	87	0.23	26	0.15	86	0.18
4	58	-0.55	34	0.41	18	-0.93	41	0.64	18	-0.84	41	0.3
5	65	-0.54	44	0.43	24	-0.35	58	0.41	24	-0.26	58	0.23
6	23	0.4	45	0.17	23	0.32	62	0.47	23	0.25	60	0.31
7	14	0	0	0	14	-0.67	43	0.4	14	-0.91	43	0.36
8	34	0.19	75	0.21	34	0.26	90	0.23	34	0.12	90	0.17
9	19	0.54	46	0.25	19	-0.36	58	0.31	19	-0.14	58	0.24
10	32	0	0	0	32	0.84	96	0.19	32	1.34	89	0.21
11	22	0.07	51	0.28	22	0.42	61	0.31	22	0.36	59	0.23

Appendix 6 Some examples of teacher assessments and overall assessments

School SCHOOL ONE **Student** STUDENT 1

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER ONE	A	5	3	3	3	5	5	3	3	5
	B	3	3	3	3	3	5	3	5	1
	C	3	3	1	3	3	3	1	1	5
	All	3 A	3 A	3 A	3 B	3 A	5 B	3 B	3 B	5 C
TEACHER THREE	A	2	5	3	3		5		5	7
	B	1	3	3	3		3		3	3
	C	1	5	3	1		1		1	5
	All	1 A	4 C	3	2 B		3 A		3 A	5 A
TEACHER FIVE	A	5	5	5			5	5	5	
	B	5	5	5			5	5	5	
	C	5	5	5			5	5		
	All	5 B	5 C	5 C			5 A	5 B	5 B	
TEACHER EIGHT	A	5	5	5		5	5	5	5	5
	B	5	5	5		5	5	5	5	5
	C	5	5	5		5	5	5	5	7
	All	5 A	5 A	5 A		5 A	5 A	5 C	5 A	6 C
Average		3.50	4.25	4.00	2.50	4.00	4.50	4.33	4.00	5.33
Standard Dev		1.91	0.96	1.15	0.71	1.41	1.00	1.15	1.15	0.58
Overall Result		-	4 - 5	3 - 5	2 - 3	3 - 5	4 - 5	4 - 5	3 - 5	5 - 6



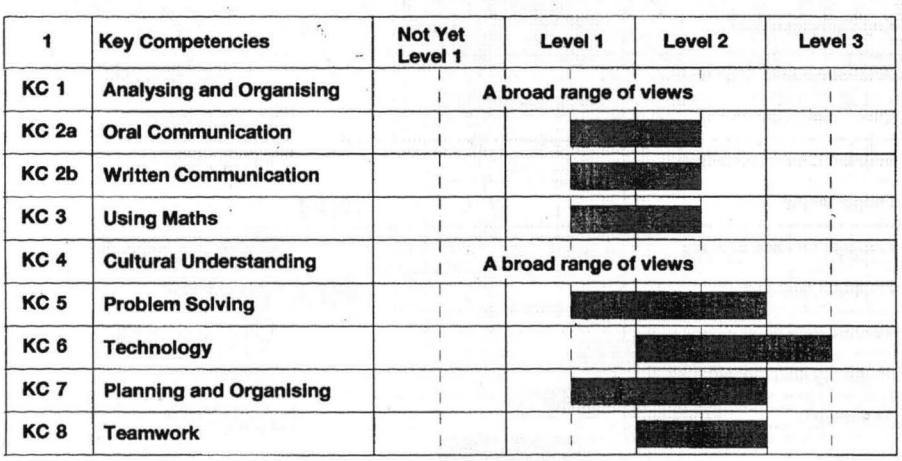
School SCHOOL ONE **Student** STUDENT 2

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER ONE	A	8	8	5		8	5	8	5	5
	B	8	7	5		8	5	7	5	8
	C	8	7	7		8	7	8	5	7
	All	8 B	7 A	5 C		8 B	5 C	8 A	5 B	7 B
TEACHER FIVE	A	8	7	7			7	7	7	
	B	8	7	7			7	7	7	
	C	8	7	7			7	8		
	All	8 A	7 B	7 B			7 C	7 C	7 B	
TEACHER SEVEN	A	6	5	6	5	5	5		7	6
	B	7	5	6	5	5	5		7	6
	C	7	5	6	5	5	5		7	6
	All	7 B	5 A	6 B	5 A	5 A	5 B		7 C	6 A
TEACHER EIGHT	A	7	7	7		7	7	5	7	7
	B	7	7	7		7	7	5	7	7
	C	7	7	7		7	7	5	7	7
	All	7 B	7 B	7		7 A	7 C	5 A	7 B	7 B
Average		7.50	6.50	6.25	5.00	6.67	6.00	6.67	6.50	6.67
Standard Dev		0.58	1.00	0.96		1.53	1.15	1.53	1.00	0.58
Overall Result		7 - 8	6 - 7	6 - 7	5 - 5	6 - 8	5 - 7	6 - 8	6 - 7	6 - 7

1	Key Competencies	Not Yet Level 1	Level 1	Level 2	Level 3
KC 1	Analysing and Organising				
KC 2a	Oral Communication				
KC 2b	Written Communication				
KC 3	Using Maths				
KC 4	Cultural Understanding				
KC 5	Problem Solving				
KC 6	Technology				
KC 7	Planning and Organising				
KC 8	Teamwork				

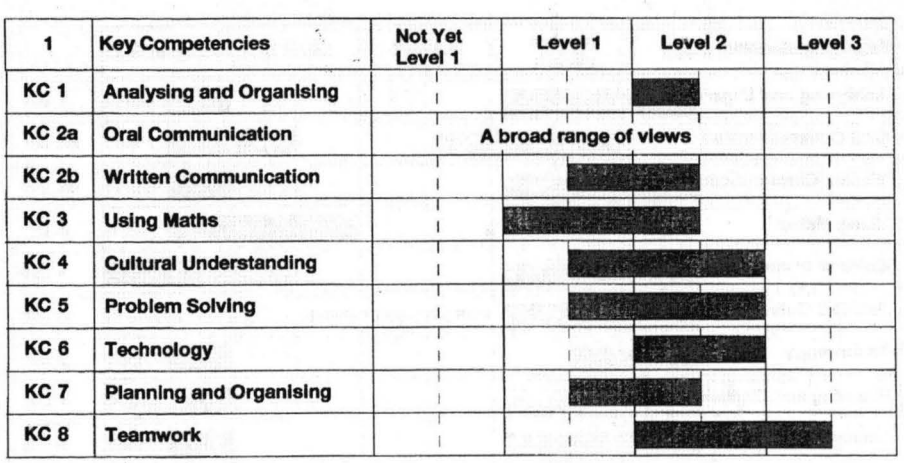
School SCHOOL ONE **Student** STUDENT 3

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER ONE	A	3	3	3	5	3	5	7	5	5
	B	3	3	3	5	1	5	7	5	3
	C	3	5	3	5	1	3	7	3	5
	All	3 A	3 C	3 C	5 A	1 A	5 A	7 C	5 B	5 C
TEACHER TWO	A	4	4	4	5		5	5	5	4
	B	4	4	4	4		4	5	4	4
	C	4	4	4	4		4	5	4	4
	All	4 A	4 A	4 A	4 B		4 B	5 C	4 B	4 C
TEACHER THREE	A	5	5	5	5		5	7	5	5
	B	5	5	5	5		5	7	5	5
	C	3	5	5	3		5	7	5	7
	All	4 B	5 C	5 B	4 B		5 A	7 A	5 B	6 C
TEACHER SEVEN	A	7	5	6	5	7	7	5	7	6
	B	8	5	6	5	7	7	5	7	6
	C	8	5	6	5	7	7	5	7	6
	All	8 C	5 A	6 B	5 B	7 B	7 B	5 B	7 C	6 B
Average		4.75	4.25	4.50	4.50	4.00	5.25	6.00	5.25	5.25
Standard Dev		2.22	0.96	1.29	0.58	4.24	1.26	1.15	1.26	0.96
Overall Result		-	4 - 5	4 - 5	4 - 5	-	4 - 6	5 - 7	4 - 6	5 - 6



School SCHOOL ONE **Student** STUDENT 4

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER TWO	A	5	5	5	5		5	5	5	5
	B	5	5	5	5		4	5	4	5
	C	4	5	5	4		4	5	4	6
	All	5 B	5	5	5		4	5	4	5
TEACHER SIX	A	5	6			6	5	7	6	8
	B	5	6			6	7	7	6	8
	C	5	6			5	7	5	6	8
	All	5 A	6 A			6 A	6 B	6 B	6 C	8 B
TEACHER SEVEN	A	5	3	3	3	3	5		3	5
	B	5	3	3	3	3	5		3	5
	C	5	4	3	3	3	1		3	5
	All	5 A	3 B	3 A	3 A	3 A	5 A		3 B	5 C
TEACHER EIGHT	A	7	7	5		7	5	5	5	7
	B	5	7	5		5	5	5	5	5
	C	5	7	5		5	5	5	5	7
	All	5 A	7 A	5 B		5 A	5 B	5 C	5 C	6 A
Average		5.00	5.25	4.33	4.00	4.67	5.00	5.33	4.50	6.00
Standard Dev		0.00	1.71	1.15	1.41	1.53	0.82	0.58	1.29	1.41
Overall Result		5 - 5	-	4 - 5	3 - 5	4 - 6	4 - 6	5 - 6	4 - 5	5 - 7



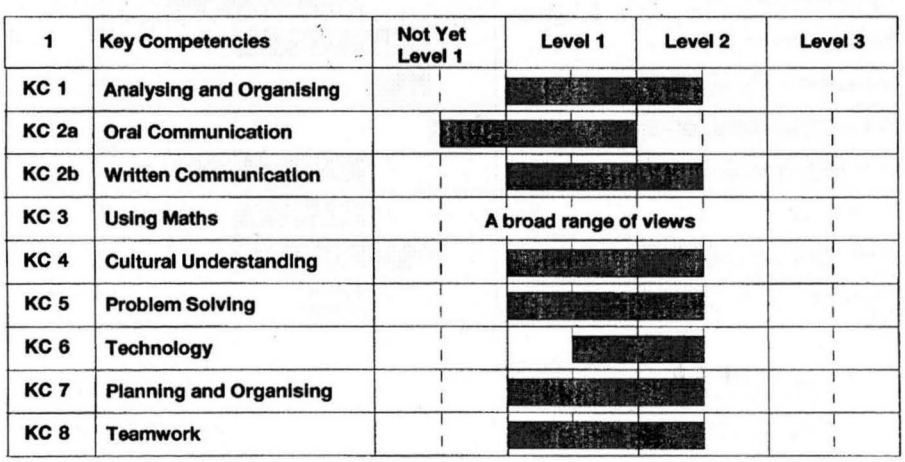
School SCHOOL ONE **Student** STUDENT 5

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER THREE	A	7	5	7	7	5	7		7	5
	B	7	5	7	7	5	7		7	5
	C	5	7	7	7		5		5	7
	All	7 A	6 C	7 A	7 C	5 A	6 A		6 B	6 C
TEACHER FOUR	A	7	7	6		8	7		7	7
	B	7	7	7		8	8		7	7
	C	6	7	6		7	7		5	7
	All	7 A	7 A	7 A		8 B	7 B		7 A	7 A
TEACHER SEVEN	A	6	6	6	5	7	3		7	7
	B	7	6	6	5	7	3		7	7
	C	7	6	6	5	7	3		7	7
	All	7 C	6 B	6 B	5 B	7 B	3 A		7 C	7 B
TEACHER EIGHT	A	7	7	7		7	7	7	7	7
	B	7	7	7		7	7	7	7	5
	C	7	7	7		7	7	7	7	7
	All	7 A	7 A	7 C		7 B	7 A	7 B	7 C	7 A
Average		7.00	6.50	6.75	6.00	6.75	5.75	.7.00	6.75	6.75
Standard Dev		0.00	0.58	0.50	1.41	1.26	1.89		0.50	0.50
Overall Result		7 - 7	6 - 7	7 - 7	5 - 7	6 - 8	-	7 - 7	7 - 7	7 - 7

1	Key Competencies	Not Yet Level 1	Level 1	Level 2	Level 3
KC 1	Analysing and Organising				
KC 2a	Oral Communication				
KC 2b	Written Communication				
KC 3	Using Maths				
KC 4	Cultural Understanding				
KC 5	Problem Solving	A broad range of views			
KC 6	Technology				
KC 7	Planning and Organising				
KC 8	Teamwork				

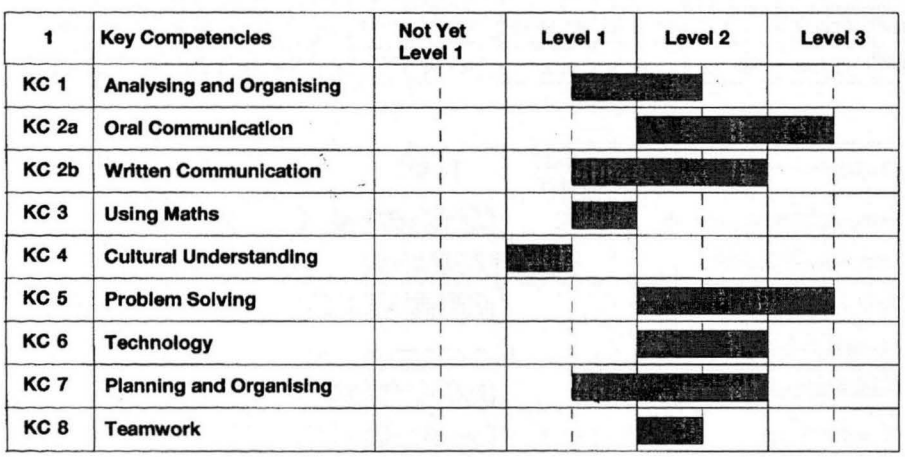
School SCHOOL ONE **Student** STUDENT 6

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER TWO	A	4	4	3	5		5	5	5	4
	B	4	4	3	4		4	5	4	4
	C	3	3	3	4		4	4	4	5
	All	4 A	4	3	4		4	5	4 B	4 B
TEACHER FIVE	A	3	4	4			6	4	5	
	B	3	3	4			5	4	4	
	C	3	4	5			5	3		
	All	3 C	4 A	4 C			5 A	4 C	5 A	
TEACHER SEVEN	A	5	1	1	1	3	3		3	4
	B	5	1	1	1	3	3		3	3
	C	5	1	1	1	3	3		3	3
	All	5 A	1 C		1 A	3 A	3 A		3 A	3 C
TEACHER EIGHT	A	3	5	5		5	3		3	5
	B	3	3	5		5	3		3	3
	C	3	3	5		5	3		3	5
	All	3 A	3 A	5 B		5 B	3 A		3 C	5 C
Average		3.75	3.00	4.00	2.50	4.00	3.75	4.50	3.75	4.00
Standard Dev		0.96	1.41	1.00	2.12	1.41	0.96	0.71	0.96	1.00
Overall Result		3 - 5	2 - 4	3 - 5	-	3 - 5	3 - 5	4 - 5	3 - 5	3 - 5



School SCHOOL ONE **Student** STUDENT 8

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER THREE	A	5	7	5	5		5		5	5
	B	3	5	5	3		5		5	3
	C	1	7	5	3		3		3	5
	All	4 A	6 A	5 A	4 A		4 A		4 B	5 C
TEACHER FIVE	A	5	7	7			5	5	6	
	B	5	5	5			7	6	5	
	C	5	7	7			7	5		
	All	5 A	7 C	7 A			7 B	5 B	5 A	
TEACHER SIX	A	4	5	4		3	7	6	6	7
	B	4	5	5		3	7	7	6	5
	C	3	5	3		3	7	6	3	4
	All	4 C	5 A	4 B		3 A	7 A	6 B	6 A	5 A
Average		4.33	6.00	5.33	4.00	3.00	6.00	5.50	5.00	5.00
Standard Dev		0.58	1.00	1.53			1.73	0.71	1.00	0.00
Overall Result		4 - 5	5 - 7	4 - 6	4 - 4	3 - 3	5 - 7	5 - 6	4 - 6	5 - 5



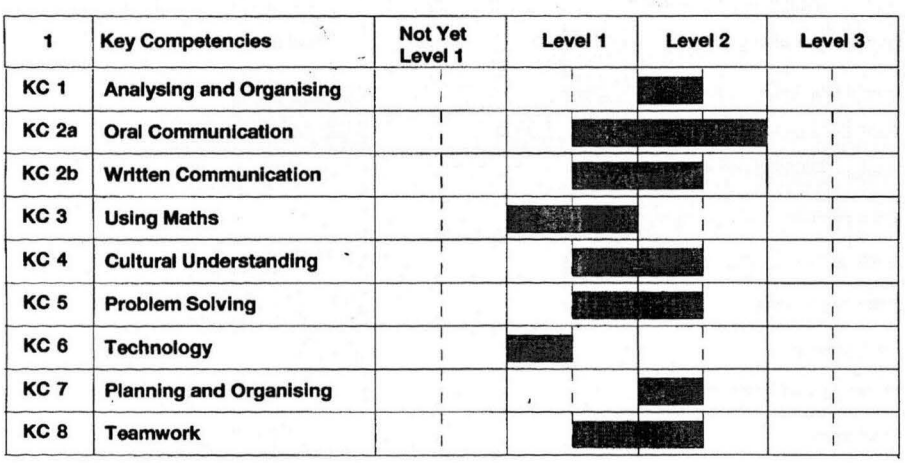
School SCHOOL ONE **Student** STUDENT 11

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER ONE	A	5	5	5		7	5	5	5	5
	B	5	5	5		5	5	5	5	5
	C	3	5	3		5	3	3	5	7
	All	5 A	5 A	5 A		5 A	5 B	5 B	5 B	5 C
TEACHER FOUR	A	7	8	7		7	7		7	7
	B	5	7	7		8	8		7	5
	C	5	7	6		6	6		5	8
	All	5 A	7 A	7 A		7 B	7 B		7 A	7 C
TEACHER SIX	A	5	7	7		6	7	7	7	7
	B	5	7	7		7	8	7	7	7
	C	5	7	6		7	7	6	7	7
	All	5 A	7 A	7 B		7 B	7 B	B	7 B	7 C
TEACHER EIGHT	A	7	5	7		5	5	5	7	7
	B	7	5	7		5	5	5	7	7
	C	7	5	7		5	5	5	7	7
	All	7 A	5 A	7 B		5 A	5 A	5 A	7 A	7 C
Average		5.50	6.00	6.50		6.00	6.00	5.00	6.50	6.50
Standard Dev		1.00	1.15	1.00		1.15	1.15	0.00	1.00	1.00
Overall Result		5 - 6	5 - 7	6 - 7	-	5 - 7	5 - 7	5 - 5	6 - 7	6 - 7

1	Key Competencies	Not Yet Level 1	Level 1	Level 2	Level 3
KC 1	Analysing and Organising				
KC 2a	Oral Communication				
KC 2b	Written Communication				
KC 3	Using Maths				
KC 4	Cultural Understanding				
KC 5	Problem Solving				
KC 6	Technology				
KC 7	Planning and Organising				
KC 8	Teamwork				

School SCHOOL ONE **Student** STUDENT 13

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER THREE	A	5	7	5	5		5		5	5
	B	5	5	5	5		5		5	3
	C	4	7	5	3		3		3	5
	All	5 A	6 A	5 B	4 B		5 A		5 B	5 C
TEACHER FOUR	A	4	5	4	4	5	3		5	5
	B	3	4	3	3	5	3		5	1
	C	3	5	3	3	3	1		4	4
	All	4 A	5 A	3 A	3 A	5 A	3 A		5 A	4 A
TEACHER SEVEN	A	5	5	4	3	3	4	4	5	7
	B	4	5	4	3	3	4	4	5	3
	C	5	5	3	3	3	4	4	3	5
	All	5 A	5 C	4 B	3 A	3 A	4 A	3 C	5 A	5 A
TEACHER EIGHT	A	5	5	5		5	5	3	5	5
	B	5	3	5		5	5	3	5	3
	C	5	5	5		5	5	3	5	5
	All	5 A	4 A	5 B		5 A	5 A	3 A	5 B	4 C
Average		4.75	5.00	4.25	3.33	4.33	4.25	3.00	5.00	4.50
Standard Dev		0.50	0.82	0.96	0.58	1.15	0.96	0.00	0.00	0.58
Overall Result		5 - 5	4 - 6	4 - 5	3 - 4	4 - 5	4 - 5	3 - 3	5 - 5	4 - 5



School SCHOOL ONE **Student** STUDENT 14

Teacher Name		KC1 Analyse and Organise	KC2a Oral Commun	KC2b Written Commun	KC3 Using Maths	KC4 Cultural Understanding	KC5 Problem Solving	KC 6 Technol Understanding	KC7 Planning and Organise	KC8 Working in Teams
TEACHER ONE	A	5	5	7	7	3	8	5	8	5
	B	5	3	5	7	3	8	7	8	3
	C	5	3	5	7	1	3	7	5	3
	All	5 A	3 A	5 A	7 B	3 A	7 A	7 C	7 B	3 A
TEACHER THREE	A	7	7	7	5		7	7	7	5
	B	7	5	7	5		7	7	7	7
	C	5	7	7	7		5		5	7
	All	6 A	7 C	7 B	6 C		7 A	7 B	7 B	6 B
TEACHER FOUR	A	8	7	7	7	7	7		7	6
	B	8	7	7	7	7	7		7	6
	C	7	7	7	7	6	5		5	5
	All	7 A	7 C	7 A	7 B	7 A	7 A		7 A	6 B
TEACHER SEVEN	A	8	6	7	5	6	7		8	8
	B	8	6	7	5	7	7		8	8
	C	8	6	7	5	7	7		8	8
	All	8 C	6 B	7 B	5 B	7 B	7 C		8 C	8 B
Average		6.50	5.75	6.50	6.25	5.67	7.00	7.00	7.25	5.75
Standard Dev		1.29	1.89	1.00	0.96	2.31	0.00	0.00	0.50	2.06
Overall Result		6 - 7	-	6 - 7	6 - 7	-	7 - 7	7 - 7	7 - 7	-

1	Key Competencies	Not Yet Level 1	Level 1	Level 2	Level 3
KC 1	Analysing and Organising				
KC 2a	Oral Communication	A broad range of views			
KC 2b	Written Communication				
KC 3	Using Maths				
KC 4	Cultural Understanding	A broad range of views			
KC 5	Problem Solving				
KC 6	Technology				
KC 7	Planning and Organising				
KC 8	Teamwork	A broad range of views			

Appendix 7 Reactions of employers to a sample report based on the Mayer Key Competencies

School Reference and Key Competencies Portfolio

An Investigation of
Some Employer Reactions
to a proposed
Sample Report

Part of the School Based Key Competencies Levels Assessment Project
funded by the Department of Employment, Education and Training

Chris Reynolds
January 1997

This Report

Following some preliminary information about purposes and methodology, this report is in four sections. The first provides an overview of employer reactions; the second a more detailed account of responses to particular elements of the Sample Report. The third section indicates the weight that may be put on these in relation to the purposes of the project and the fourth summarises the findings and offers some recommendations.

1. Preliminary

1.1 Purpose and some assumptions

The purpose of the investigation is to elicit employer reactions to a Sample Report developed through the project. A national perspective on the reactions of employers is required because the Sample Report relates to a scheme which would entail employers across the country being able to make effective use of the information it will convey.

Given that some groups, especially those in larger businesses where specialists in the Human Relations and Recruitment fields are available, are likely to be more familiar with the Key Competencies and the background to their development, it is important to seek reactions from such organisations. It is equally important to seek the reactions of medium and small business people amongst whom awareness of the Key Competencies is less widespread and less in-depth.

These two groups have been identified and canvassed using different strategies. Similarly, both quantifiable and qualitative responses have been sought.

1.2 Methodology

To these ends both questionnaire/survey and telephone interview methods have been employed (see Appendices 11). The survey was designed to gather a wider range of responses and interviews to provide specific and more detailed comments on certain aspects of the Sample Report. The interview was designed to confirm certain responses and to provide opportunity to probe and elaborate others - in particular to elicit suggestions for improving the Sample Report in Part 5 pages 61-64.

1.3 The Survey

The survey solicited responses to elements of the Sample Report and to questions of assessment reliability. It also sought views on the qualities which employers seek in young job applicants as well as what other information they would like access to.

The survey was circulated to 430 business addresses. To identify these two mailing lists were used. The first was supplied by the National Industry Education Forum and included large

businesses and a number of peak employer organisations. It contained some 180 names. Of these 32 (18.3%) responded. The second was a list of 250 businesses arbitrarily selected from the National Yellow Pages [11]. From this group there were 5 survey responses and 9 interviews. A telephone follow-up was conducted with a sizable proportion of those on this second list who had not responded. The follow-up revealed a range of reasons for non-response. They seem to reflect a sense in which small (one person) businesses do not see school exit documentation as relevant because they rarely employ young people and, when they do, they use other more personal techniques. This information can itself be used constructively to argue broadly for modifications to the report format. More detail and some interpretation of the strength of the data with further commentary on the low response rates can be found in Section 3.

Table 1 summarises the survey respondents.

Table 1 Survey Respondents by state/territory and type of business

State	Consultancy	Corporate	Peak	Service	Small Professional	Small Trade	Totals
ACT	0	1	1	1	0	0	3
NSW	0	6	2	2	3	0	13
NT	0	0	0	0	0	1	1
QLD	0	0	1	0	1	0	2
SA	1	1	0	0	0	0	2
Vic	1	12	2	0	0	0	15
WA	0	2	1	0	0	0	3
Totals	2	22	7	3	4	1	39

1.4 The Interview

Survey respondents were asked whether they would also be prepared to take part in a short telephone interview. Of the 39 survey respondents 26 also took part in interview. In addition, before surveys had been returned, some potential respondents were approached for interview. There were 9 such interviews. Table 2 summarises the interview respondents.

[11] The distribution of businesses identified was designed to match state population figures but with a weighting towards the smaller states and territories. The weighting was to balance the fact that large businesses are commonly have headquarters in the metropolitan centres of one or other of the larger states.

Table 2 Interview Respondents by state/territory and type of business

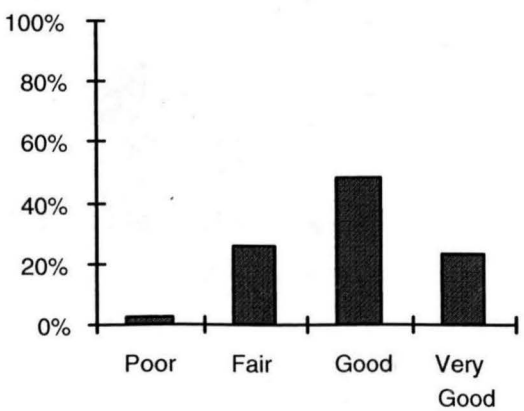
State	Consultancy	Corporate	Peak	Service	Small Professional	Small Trade	Totals
ACT	0	2	2	0	1	0	5
NSW	5	1	2	7	0	0	15
NT	0	0	0	0	1	0	1
Qld	0	1	0	0	0	0	1
SA	0	0	0	0	0	1	1
Vic	5	2	0	0	0	1	8
WA	2	2	0	0	0	0	4
Totals	12	8	4	7	2	2	35

Section A

Overall Outcomes

To establish the assumptions being made respondents were asked how they would rate the Key Competencies as a basis for reporting about student performance to employers. Figure 1 summarises the responses.

Figure 1 How would you rate the Key Competencies?



Only one respondent was negative about the Key Competencies and equally negative about the Sample Report. Reasons given by this respondent were associated with conceptual difficulties relating to standards

What information, in addition to that offered by the Key Competencies, would be most valuable?

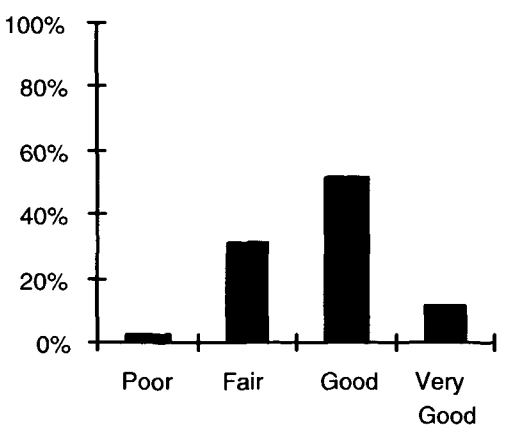
Responses indicated that a range of additional information would be required. This included information about any prior work and in the case of school leavers of work experience. In some instances assessments of performance in specific subject areas would be required.

Finding 1

Respondents found the Key Competencies to be a good basis for schools reporting student performance to employers but they require other information as well.

Towards the end of the survey respondents were asked to rate the kind of information provided in the Sample Report (Question 14). Figure 2 summarises the responses to this question.

Figure 2 How would you rate the Sample Report?



All but one of the responses were positive. Some of the detailed commentary which accompanied the single negative response was substantive, but most was based on misreadings of the Facet descriptors and their relationship to each other. These matters are taken up in the next section.

As might be expected, the correlation between responses to questions 13 and 14 was strong (r=0.63). That is, there was strong support for the Key Competencies and more modest but still strong support for the Sample Report presented.

In general this and the interview data show that respondents warmed to the Sample Report. However, many found it too complex and not as easily understood as they required. This was especially true about less familiar elements such as the graphic display [element 1] and the report of the Facets [element 2]. These elements had been included, it will be appreciated, to provide assessments of levels of achievement. While there was general agreement that levels of achievement information was to be welcomed (only 4 respondents said "No"), there was a range of views about how much confidence could be placed in the information.

Finding 2

Respondents find the information provided in the Sample Report to be good, but for most its presentation was too complex and unclear. This is particularly the case with elements 1 and 2.

Section B Detailed Outcomes

What employers look for in selecting juniors and school leavers for employment

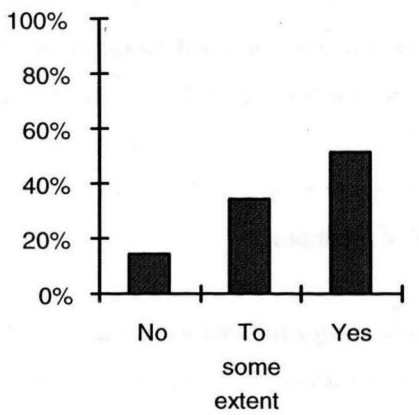
What employers look for varies, but it can be clustered around a set of 'soft skills' or personal attributes. The range of such attributes described by respondents varies between personal presentation to a capacity to 'solve problems', between being able to communicate and to work effectively with others. Corroborating the findings of others (Stanton 1995, DEETYA 1996), the degree of agreement between these statements and the Key Competencies themselves is very high. In addition to the assessments of Key Competencies, respondents look for other information - ranging from academic results to information about work experience and career interests. When responses (question 16) are broken down by business type there is a tendency for Corporate Human Relations managers to prefer detailed and varied information about different aspects of student achievement and for those in small business to seek something which conforms with their own experience of educational assessments usually limited to school references and reports with grades and marks. Two interviewees expressed concern that the Key Competency assessments would replace subject assessments and, when reassured that this was not intended, were then very positive about the role Key Competency assessment could play to complement subject specific reporting.

Element by Element Responses

Element 1: Reporting Levels of Performance

Survey respondents were asked if, in looking for employees, summary levels would be informative and useful. Figure 3 reports an analysis of their responses.

Figure 3 Would summary levels of performance be useful?



A majority of respondents welcomed the assessment of levels of performance in the Key Competencies and were positive about the succinctness of the information provided in the graphic display, its capacity to allow for comparisons between applicants and the capacity for quickly ascertaining the overall pattern of an applicant's strengths and weaknesses. A number of respondents specifically welcomed the proposal that there should be a common national format, because this would prima facie provide a more effective basis for comparing job applicants.

Reservations were expressed by 5 respondents. Of these 2 were concerned with the usefulness of the level of performance information ("comparative measures tend to reflect irrelevant variables - SES etc", "Bands of ratings - did not feel it served any purpose."). Other reservations were relatively simple matters of form and presentation which are discussed elsewhere in this report. They are not however, unimportant. Several responses were similar to the one which referred to the Sample Report as not being "management friendly" or "not an educators' document".

Only 15.2% of respondents said that summary levels of achievement would not be useful. Of these one worked with a major corporation, another two with peak national employer bodies and the fourth as a small trade business. The peak employer rarely hires school leavers, the small businessperson selects using word of mouth and school reports (ie grades and

comments). All four expressed preference for marks and grades in school subjects. All four indicated that it is their practice to check references and/or to use word of mouth information about an applicant's strengths and weaknesses. All five of these rated the Key Competencies as a fair (3:5) or good (2:5) basis for school reporting about students to employers. However, their responses were less enthusiastic than were those for all respondents (Mean=2.36).

Other respondents supported the notion of levels of Key Competency performance being reported, although some were confused about the nature of the levels. Two respondents were not sure about the direction of the scale ("one presumes level 3 is the more advanced level") and the same respondent could not anchor the scale in any known points ("this document presupposes a standard but does not define the standard ie is the standard based on the level of education attained by the students or is it based on the requirements of an employer?"). These two issues need to be attended to both substantively and by providing on the reporting format clear information to users about what it is they are looking at.

There was strong support for the attempt to report levels of performance and, with two exceptions, a preparedness to accept the judgments provided as valid and useful. The exceptions were one which expressed concern about the potential for the bar to be read from the bottom rather than as an indication of a range (perhaps a box and whisker may be better suited to conveying the notion of the boundaries of agreement and the likelihood of the agreed assessment) and those that felt the need for the levels to be more clearly defined (ie what tasks have been completed which illustrate the level of performance achieved). Here it may be useful to look at the possibility of linking level assessments with the particular details provided elsewhere in the report format (especially on page 3). This would enable a quick appraisal to be made of what a student has done which may illustrate why such a level was corporately determined by teachers. However, this would need to be done in a way which does not imply that the details were the basis for the corporate judgments.

Finding 3

Levels of performance information is welcomed by almost all respondents. The concerns of those who do not welcome are focused largely express on presentation and the lack of clear information about the meaning of what is before them.

Clearer specification of the standards communicated by the scale and clearer communication of the direction of the scale are required.

Element 2 - School Facet highlighting or specific comment

Figures 4 and 5 report responses to the notion of Facets and the proposal that a standard Facet comment could be replaced with a customised distinctive comment where this would be appropriate.

Figure 4 Reporting of a strong or typical Facet useful?

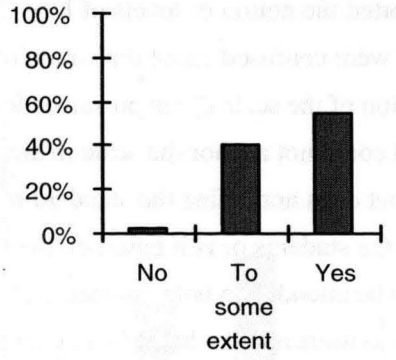
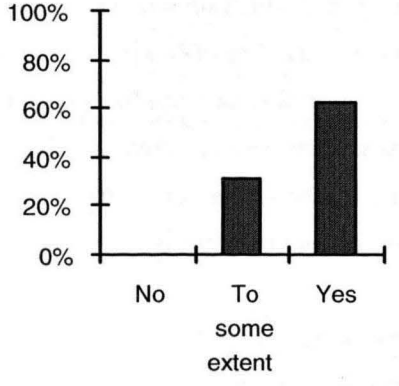


Figure 5 Reporting of a distinctive comment useful?



Figures 4 and 5 indicate strong support for the Facets and this despite of the frequency of critical comments both in the survey and in interviews.

As we have seen, there were some detailed comments about the format used to report element 2 as well as about the concepts underpinning the element. Few of these were concerned with substantive issues. No respondent suggested the need to reconceive of the major categories represented in the Key Competencies themselves and only a few comments

were made suggesting the need to modify the way in which these were broken down into Facets.

Yet this element, as presented in the Sample Report, was seen most frequently to be problematic. That is, although respondents clearly signalled an interest in detailed information (one respondent expressed interest in differing aspects of *working with others and in teams*[12]), some reported that the pattern of Facets was not adequate to match the range required in workplace settings.
Repeatedly discomfort was expressed with the way information in section 2 was presented. Two respondents interpreted the use of the A/B/C code to indicate a hierarchy of levels of performance and then reported their disagreement with how the Facets were constructed. Others did not detect the difference between a school Facet comment and a specific one. Still others found the repetition of the numbers (cols 1 and 3) to be confusing. Another felt that the information in this section could have been built into the reporting in element 1 but did not specify how this might be done.

One detailed comment raised the issue of the conceptual links between the Summary (element 1), the Facets (element 2) and the School Statement. At the moment the format does not clearly link these elements so that each can be seen to reinforce, explicate or to summarise each the others. There may need to be conceptual refinement in this regard.

Finding 4 *The concept of Facets and the descriptors provided were widely welcomed by respondents and the concept of distinctive comments was also strongly supported. However, important improvements are required.*

In summary three improvements need to be made.

1. Render the information in as clear and as graphic a fashion as possible. For example, symbols rather than numeric codes could be employed; unique comments could be highlighted to distinguish them from standard descriptors.

2. Provide explications (they would need to be brief) of the Key Comptency scale, the Facets (rather than having to turn a page) and the relationship between each of

[12] See the work of Stasz and others which suggests that the generic competencies have quite different emphases in different workplace settings. That is self-managing teams may be required in one context and knowledge/authority teams (with clear role differentiation) in another. Stasz C., Ramsey K., Eden R., Melamid E. and Koganoff T., 1996 *Workplace Skills in Practice*, Santa Monica CA, Rand/National Center for Research in Vocational Education

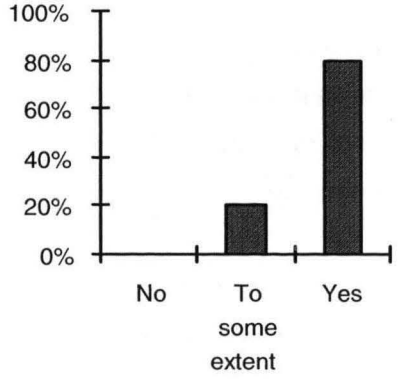
the elements of the school's assessment. These requirements may present significant but not insuperable challenges to layout-design and copywriting.

3. Render the relationship between the summary global judgments, the Facet descriptors/distinctive comments and the continuous prose of the School Statement transparent to the user.

Element 3

This element was familiar to most respondents since it is the one most closely resembling the type of school reference schools currently issue. Figure 6 reports responses to question 5.

Figure 6 Do you think the school statement would be informative and useful?



Few respondents express any reservations at all about it and approximately 80% stated that it would be useful and informative.

Finding 5 *The School Statement is very strongly supported and should be seen to explicate and elaborate the summary information which precedes it in the Sample Report.*

Monitoring school assessments - a necessary digression

Responses to questions 6, 7 and 8 relate to the issue of monitoring. Respondents were asked to indicate whether assessments would be sufficiently reliable. Figures 7, 8 and 9 report responses.

Figure 7 Would assessments be sufficiently reliable without external monitoring?

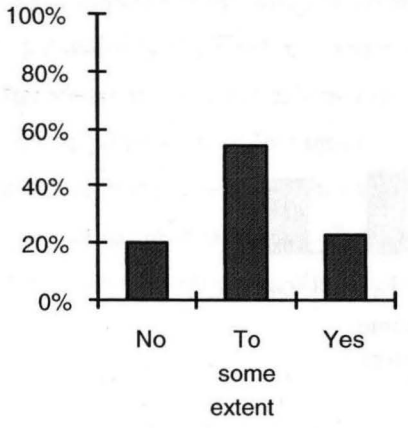
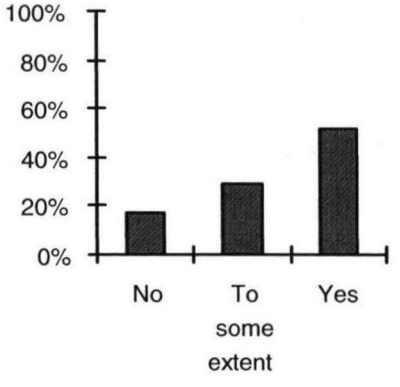
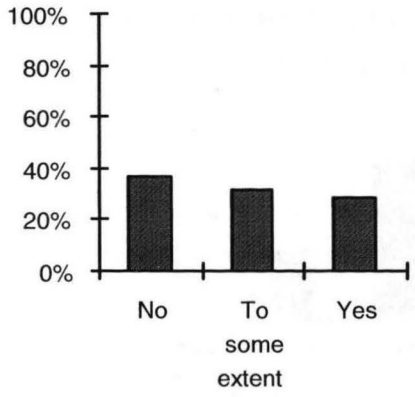


Figure 8 **Would the external monitoring of assessments made by schools make them more useful?**



As a further check respondents were asked whether their answers to earlier questions about school statements would have been more positive if the school assessments had been subject to external monitoring by a Board of Studies or other external agency. Figure 9 reports the responses.

Figure 9 **Would your response be more positive to earlier questions if school assessments were externally monitored?**



It may be deduced from the distribution shown in Figure 7 that a sizable proportion of respondents, having already responded positively did not see that external monitoring would add very much. The correlations between responses to these three questions and between those for questions 1-4 are reported in the following table.

Table 2: Inter-Item correlations, questions 1, 3, 4, 5, 6, 7 and 8

Item	1	3	4	5	6	7
3	0.39					
4	0.02	0.43				
5	0.01	-0.08	0.05			
6	0.07	-0.14	-0.27	-0.20		
7	0.17	0.21	0.21	0.22	-0.26	
8	0.15	-0.10	0.08	0.15	-0.05	0.57

Correlation coefficients reported in Table 2 consistently support the view that respondents would value external monitoring. That is the positive correlation between responses to questions 7 and 8, both for the use of monitoring are consistent with the negative correlations between these responses and responses to question 6.

The relatively low positive correlations between responses to question 1 and the questions 6, 7 and 8 on monitoring are not strong enough to suggest any consistent view although the tendency is towards external monitoring. Similarly, the relationship between responses to questions 3 and 8 may in part be due to the fact that those supporting Facets believed that they could not be more supportive than they had already indicated they were. Given these kinds of reservations, further probing of these relationships remains problematic.

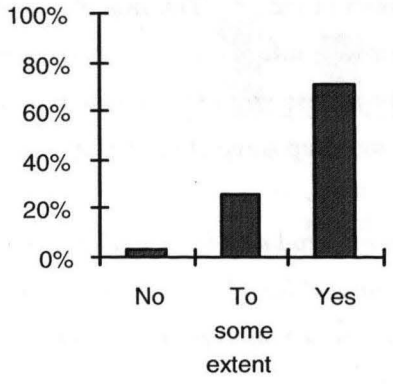
The major substantive question is whether employers can come to accept that the information provided is to be relied on. The issues of inter-school comparability and the related question of the fairness to students of the way performance assessments are produced was raised. As we have seen, a majority (82%) of respondents say that their confidence would be increased if the assessments were subject to external monitoring. But here there was also some ambiguity - attributable it seems to a number of factors. The first seems to be a reluctance among some to credit the levels of performance information; a second, among others, was a preference for qualitative information. Both these groups are small as a proportion of the whole set of responses, but together they make up some 20% of responses.

Finding 6 *The monitoring of the assessments produced by schools is seen to be highly desirable by the great majority of repondents, although some were so strongly in support of the Sample Format that they felt they could not be more positive than they were.*

Element 4

Figure 7 reports responses to the proposal for a Student Statement as presented in the Sample Report.

Figure 10 Do you think a Statement by the students about his or her performance on the Key Competencies would be useful and informative?



Element 4 also was familiar to respondents and was regarded favourably because it provides an employer with the opportunity to consider the applicant in their own voice. Employers who are looking to get to know him or her find this material most likely to give them a sense of the applicant's personality in ways which the quantified information provided elsewhere, they say, does not.

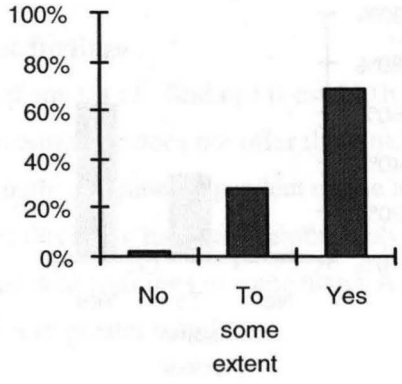
Concern was voiced about the extent to which the text would be produced by the student her/himself; that professionals or teachers in the school might "over-assist" the student and thus vitiate the value of the information and call in question the integrity of all such material. Ways to address this concern need to be found.

Finding 7 *There is clear support for the inclusion of a Student Statement, but concerns voiced about how it can be authenticated and therefore equitably regarded. The issue is to be resolved as to whether the school does this or whether the student is regarded as a free agent in assembling this and other material.*

Element 5

In the Sample Report a student provides a list of specific skills any or all of which may relate to the job being sought. Responses to question 10 are reported in Figure 11.

Figure 11 Would a statement of Student Specific Skills be informative and useful?



The list of specific skills provided in this element was warmly welcomed in most responses, but concern was voiced about the relevance of the list to specific jobs and job seeking processes. There are simple solutions to this problem - namely that the document can be prepared in different versions each of which is tailored to the specific job-seeking initiative a student might embark on.

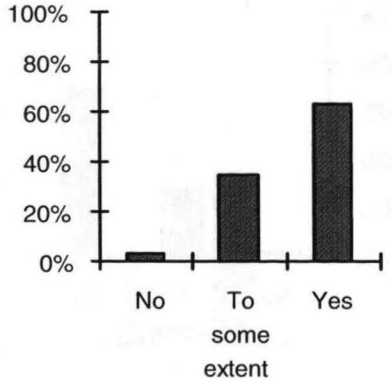
Of greater difficulty is to ensure that the skill levels indicated be accurate and comparable. Such generic statements as "has used a word-processing package" are not sufficiently precise to be useful, whereas "has been awarded a Bronze Certificate for Swimming" is because there are publicly declared standards applying to the achievement.

Finding 8 *There is general support for the inclusion of Specific Skills in the list provided by the student. This list will be more useful if the information included in it is precise and certificated or authenticated. It will be in the student's interest to ensure that this is the case.*

Element 6

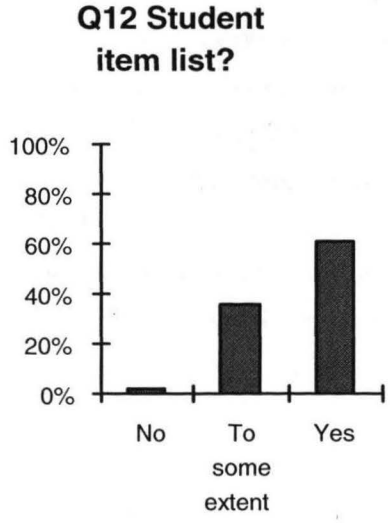
This listing is a selection from a bank of material which the students may have accumulated over a number of years. The recentness of each item and its relevance to the job application in question would be an important criterion for inclusion. Another would need to be that items listed should be relevant to a particular job-seeking purpose. Figures 12 and 13 summarise responses to questions regarding the utility of this information.

Figure 12 Would examples of student work in a portfolio be useful?



Respondents by and large welcomed the concept and recognised that access to the material was an option and one likely only to be exercised with applicants who reached an interview.

Figure 13 Do you think this index is useful?



Respondents by and large welcomed the concept and recognised that accessing the material was an option they would usually exercise only with applicants selected for interview. The recentness of each item and its relevance to the job application in question would be an important criterion for its inclusion. Another would need to be that each item should be relevant to a particular job-seeking purpose.

Finding 9 *There is general support for the inclusion of a list of portfolio items especially if the selection is targetted to a particular job application and related to the Key Competencies.*

Section C The strength of the findings

The measure of the strength of any set of findings is established by the breadth of the database established. This investigation does not offer that kind of strength. Rather it provides another kind of strength; one more dependent on the level of expertise of the respondents. As we have seen, this study has drawn extensively on the views of those in senior human relations management positions in some major Australian companies. It is useful to consider these matters in greater detail.

Response Rate

Overall less than 10% of those invited to participate responded [ie NIEF list 18.3% and Yellow Pages <5%]. The low response rate is attributable to a perception, repeatedly expressed when businesses were approached, that the information did not seem relevant to their situation. This was despite the fact that, when asked what attributes were desirable in a job applicant, respondents frequently provided a list which closely resembled the Key Competencies.

Just over 100 follow-up phone calls were made and only in four cases did the response lead to undertakings to send the survey response back. None of these has been received as of the date of this report. Verbal responses indicated 'not interested', 'not applicable' 'head office handles all this' or 'too busy'. It appears that a large proportion of those contacted had not or were not likely in the immediate future to employ a young person or anyone else for that matter.

This pattern of reactions may be used to support the argument that the distributions of survey responses and the range of responses and interpretations provided in interview represent an overly positive reflection of employer receptivity to the sample format. An equally persuasive argument is that methods for identifying appropriate respondents were quite inadequate to the task and that not too much should therefore be inferred from the low response rate. If the first argument is accepted, then the lack of response from small business people should be interpreted as rejection of the Sample Report as irrelevant to their needs.

This means that the reservations expressed by those who *did* respond will need to be taken all the more seriously if whatever scheme will be adopted is to be credible with a wide enough cross-section of employers to make its implementation worthwhile.

Neither set of responses can be argued to be "employer views" in any representative sense as the problem of identifying an appropriate population, much less to target sub-populations and therefore of creating a large enough sample exceeded the time and resources available. However, when taken together with the findings of other research (ie Stanton 1995, DEETYA 1996), the findings of this work are not inconsistent with those of larger studies.

The Repondents

With the exception of the small number of respondents who were from the small business sector, the majority came from those who were already familiar with the Key Competencies, had experience of and expertise in recruitment, training and other aspects of Human Relations management in large companies. These respondents also include those who have previously indicated that they see performance appraisal and other personnel management strategies as being closely related to the use of Key Competencies. In other words there is a significant group of senior people in business and industry who are supportive of the Sample Report and are keen to use made of it by schools and employers in employment related decision-making.

In summary, this investigation constitutes an in-depth national consultation with some small, medium and large employers from a range of industries. Its findings and recommendations should be seen in this light. While responses were favourable overall, there are clear messages from potential users (employers) as to what improvements are required.

Section D Some Findings
This section provides a summary the findings of this investigation.

Finding 1

Respondents find the Key Competencies to be a good basis for schools reporting student performance to employers.

Finding 2

Respondents find the information provided in the Sample Report to be good.

Finding 3

Levels of performance information is welcomed by almost all respondents. The concerns of those who do not welcome are focused largely express on presentation and the lack of clear information about the meaning of what is before them.

Clearer specification of the standards communicated by the scale and clearer communication of the direction of the scale are required.

Finding 4

The concept of Facets and the descriptors provided were widely welcomed by respondents and the concept of distinctive comments was also strongly supported. However, important improvements are required.

Finding 5

The School Statement is very strongly supported and should be seen to explicate and elaborate the summary information which precedes it in the Sample Report.

Finding 6
The great majority of respondents see the monitoring of the assessments produced by schools to be highly desirable, although some were so strongly in support of the Sample Format that they felt they could not be more positive than they were.

Finding 7
The is clear support for the inclusion of a Student Statement, but concerns voiced about how it could be authenticated and therefore equitably regarded. The issue is to be resolved as to whether the school does this or whether the student is regarded as a free agent in assembling this and other material.

Finding 8
There is general support for the inclusion of Specific Skills in the list provided by the student. This list will be more useful if the information included in it is precise and certificated or authenticated. It will be in the student's interests to see that this is the case.

Finding 8
There is general support for the inclusion of a list of portfolio items especially if the selection is targetted to a particular job application and related to the Key Competencies.

References

Department of Employment, Education and Training, 1996,
Benchmarking Study # 95/07/1381, DEET, Canberra

Stanton, J., 1995, Business and Industry and Key Competencies, Report of a Project,
National Industry Education Forum, Melbourne

Stasz, C., Ramsey, K., Melamid, E., and Koganoff, T., 1996, Workplace Skills in Practice,
Santa Monica CA, Rand/National Center for Research in Vocational Education.

Appendix 11 A Survey of employer reactions to school based Key Competencies Assessment and Reporting

A Survey of Employer Reactions to School-based Key Competencies Assessment and Reporting

The aim of this research is to gather reactions from employers to some ways schools might offer information to those seeking employees about the performance of students who are leaving school.

The proposals to be considered are organised around the Key Competencies. These competencies were developed by the Mayer Committee (including representatives of business, government and education) to identify the most important abilities young people will need to cope with work in the future. The Key Competencies are said to be important for work, education and life.

	THE KEY COMPETENCIES
1	Collecting, analysing and organising ideas and information
2	Communicating ideas and information
3	Planning and organising activities
4	Using mathematical ideas and techniques
5	Solving problems
6	Using technology
7	Working with others and in teams
8	Cultural understanding

The Key Competencies are outlined in more detail on Page 4 of the Sample Report. Please note that each Key Competency has 3 Facets which are the basis of the assessment, and note that one Facet may be identified in Part 2 of the report as the particular strength or most typical characteristic of the student.

Please look at the Sample Report, particularly pages 2 and 3. We wish to get respondents reactions to each part of this Report. We wish to know if employers feel that the information contained in each part of the report would be useful and informative in making employment decisions.

Please note that Section One is the assessment of a school taking into account a range of teacher judgements. Assume that these judgements are produced by the school without any external monitoring of the school decisions by a Board of Studies or other agency. Section Two is the self-assessment of the student.

Please leave aside any concerns about particular comments in the example, and respond in general terms to the kinds of information it offers. The following questions aim to gather reaction to the parts of the Sample Report.

Respondent's name: .....

Organisation: .....

Phone: ..... **Fax:** .....

Please review the Sample Report and then circle the alternative that is closest to your view of the kind of information in each part of it. You are invited to write answers to questions 15 and16.

Section One School Statements

Part 1 School Summary levels
This graphical display shows an 8 point scale for each Key Competency. The performance of students is indicated by a single position or a band of up to 3 positions on the scale. These levels are based on the judgements of a number of teachers.

1 In looking for employees, do you think such summary levels would be informative and useful?

No to some extent Yes

2 Are there too few or too many levels?

Too few about right Too many

Part 2 School Facet Descriptors
In this section of the Report, 1 of the 3 Facets of each Key Competency (indicated by the bracketed code) or a comment (without bracketed codes) is offered to sketch the strongest or most typical feature or the student's performance.

3 Do you think the reporting of a strong or typical Facet would be informative and useful?

No to some extent Yes

4 Do you think the reporting of a distinctive comment would be informative and useful?

No to some extent Yes

Part 3 School Statement
This is a statement by the school about the specific performance of the student on particular aspects of the Key Competencies. It draws on specific comments from individual teachers and is about particular situations and subject areas.

5 Do you think the school statement would be informative and useful?

No to some extent Yes

You were asked to assume in the responses above that the assessments were produced by the school without external monitoring by a Board of Studies or other agency.

6 Do you feel, in general, that assessments produced by schools without external monitoring would be sufficiently reliable to be informative and useful?

No to some extent Yes

7 Do you feel that external monitoring of the assessments made by schools would make them more useful?

No to some extent Yes

8 Would your responses to questions 1 to 4 be more positive if the school assessments were subject to external monitoring by a Board of Studies or other agency?

No to some extent Yes

Section Two Student Statements

Part 4 Student Key Competencies statement
This is a statement written by the student about her or his performance on the Key Competencies.

9 Do you think such a student statement would be informative and useful?

No to some extent Yes

Part 5 Student statement of specific skills
This is a list of specific skills a student claims to be able to demonstrate.

10 Do you think such a student statement of specific skills would be informative and useful?

No to some extent Yes

Part 6 Student portfolio items
This is a list of items in the Stidemt Portfolio and which could be consulted as evidence of the nature and quality of the student's work.

11 Do you think the examples of student work in the portfolio would be informative and useful?

No to some extent Yes

12 Do you think this index is a useful way of presenting the examples of student work in the portfolio?

No to some extent Yes

13 How would you rate the Key Competencies as a basis for schools reporting about student performance to employers?

Poor	Fair	Good	Very Good

14 Overall, how would you rate the kind of information in the Sample Report?

Poor	Fair	Good	Very Good

15 In selecting juniors and school leavers for employment, what are the main things you are looking for in them?

.....

.....

.....

.....

.....

.....

16 What information, in addition to that offered by the Key Competencies, would be most valuable for schools to offer to employers about school leavers?

.....

.....

.....

.....

Would you be prepared to accept a 10 minute phone call from us, at a time that suited you, so you can further explain your views on these matters? If so, please nominate two dates and times when we might ring you.

Date 1: ..... Time 1: .....

Date 2: ..... Time 2: .....

Thank you for your participation in this survey.