USING CASES FOR SCHOOL PRINCIPAL PERFORMANCE STANDARDS: AUSTRALIAN AND NEW ZEALAND EXPERIENCES

HELEN WILDY[1], WILLIAM LOUDEN[2] AND JAN ROBERTSON[3]

1. *Edith Cowan University, Western Australia*
2. *Edith Cowan University, Western Australia*
3. *University of Waikato*

ABSTRACT *This article reports an international validation of a framework for performance standards for school principals. The framework, generated in Australia in 1996-1997, was applied in New Zealand in 2000. The framework involved an innovative method of establishing standards for principals' performance, based on a combination of qualitative and quantitative methods of making judgements about the quality of principals' work. Using cases of critical incidents in which principals made decisions in their everyday work, the essential elements of quality performance, together with a set of dimensions on which performance varied, were identified. This study explored the application of these essential elements and dimensions in a context in which similar school system restructuring is in progress. Three questions were addressed: How valid are the Australian cases in the New Zealand setting? How applicable to the New Zealand setting are the three continua—duties, interpersonal skills, moral dispositions—that comprise the framework? And finally, How applicable to the New Zealand setting are the particular duties, interpersonal skills and moral dispositions? This study supports three conclusions. First, cases generated in one context are not applicable in a different cultural setting. However, the method of developing cases is readily applied cross culturally. Second, the values underpinning the framework developed in Australia are similar to, but not the same as, those about which principals in New Zealand assess principal performance. Third, there are similarities, but also important subtle differences, in the particular dimensions on which the framework is grounded. The study indicates the validity of using cases to generate performance standards for school principals.*

INTRODUCTION

Superficially, educational leadership in Australia and New Zealand ought not be very different. We are both former British colonies; our flags are similar enough to be confused in Olympic medal ceremonies; our accents seem the same to North American ears; we have taken yachting's highest prize (the America's Cup) from older, richer and larger countries; our rugby union teams both expect to come first in the World Cup; our netball teams vie for supremacy, and our students do equally well in international comparisons of school science, mathematics and literacy learning. The educational differences between the two countries, however, are masked by these historical, geographical and social similarities. New Zealand is a small and predominantly rural country with a radically decentralised government school system. Most of the schools are very small and all of them are managed by locally elected school boards. Australia is a much larger and more urbanised country, with most government schools centrally controlled by one or

other of the six state government school systems. The impact on school leadership of these two systems is illustrated in this paper, which concerns the development of professional standards for school principals in these two countries.

The notion of "standards" is a central metaphor of educational reform. It has been applied in a wide range of contexts, including curriculum content, students' performance, teachers' performance and principals' performance. For example, the notion of standards is at the heart of school principals' performance assessment processes in the United States (Council of Chief State School Officers, 1996a), England and Wales (Teacher Training Agency, 1996), New Zealand (Ministry of Education, 1998) and Australia (Queensland Department of Education, 1996). There is a strong similarity among contemporary attempts to develop standards for school principals. Each identifies a number of key responsibility areas such as educational leadership, staff management, and financial management. Each of these key areas is subdivided into a further number of key competencies such as "develop a shared vision for the school" and "provide educational leadership to the school and community". What these approaches have in common is that they comprise long lists of duties and dispositions. One strength of many such lists is that the development processes are often hugely consultative. There have been, for example, 10,000 school administrators involved in the development of the Interstate School Leaders Licensure Consortium Standards (Council of Chief State School Officers, 1996b). A second strength of these approaches is that they provide highly elaborated maps of the territory of principals' work.

However, setting standards through long lists of duties and dispositions has serious weaknesses (Louden, in press). Among these weaknesses are the apparent fragmentation of the work, the tendency toward dichotomous judgements, and the absence of contextual details. An alternative approach to the development of standards framework uses cases, brief narrative accounts depicting principals facing tensions and dilemmas and solving problems in their daily work. We argue that cases supply circumstances and timing (Louden & Wildy, 1999a), the contextual details that characterise complex and uncertain professional environments. A case-based standards framework for principals' performance was developed in Australia during 1996-1997. Establishing the framework involved an innovative combination of qualitative and quantitative methods of making judgements about the quality of principals' work. Rather than backgrounding context and assuming that the variation between good and bad performance can be accounted for in an undisclosed scoring rubric, this alternative case-based method makes explicit both context and quality.

In addition to its innovative use of cases, this standards framework has three unique features. First, judgements are made about principals' performance not only in terms of what they do, their duties, but also in terms of how they do their work – their interpersonal skills and moral dispositions. While sets of duties, such as managing staff, managing the curriculum and so on, are familiar enough to most practitioners, the sets of interpersonal skills and moral dispositions derived inductively from the pilot phases of the research are less familiar. The interpersonal skills identified are listening; collaborating; leading; negotiating; and confronting complexity. The moral dispositions identified are flexibility and creativity; courage and decisiveness; fairness and consistency; sensitivity and tact; and patience and persistence.

Second, the framework is unique because it captures what school administrators consider the essence of high quality performance. The essence of high quality performance is captured in a set of descriptors generated by the 1,000

school administrators who rated a set of 74 cases. Of almost 4,000 descriptive words or phrases, more than one third were accounted for by only six qualities: being supportive, decisive, fair, innovative, collaborative, and flexible.

Third, the framework is unique because it captures how school administrators consider the quality of performance varies. The variation of performance represented by each case is indicated by its position on each of three continua representing duties, interpersonal skills, and moral dispositions. In these ways, the framework uses cases in an innovative manner to generate and represent both what is seen by school administrators in Australia to be essential about principals' performance and also to demonstrate variation in quality. Further details of the method and framework appear in publications (Louden & Wildy, 1999b; Wildy & Louden, 2000).

This study sought to ascertain whether the method of developing such a standards framework was also applicable in an international context. Is the framework grounded in a set of concepts and values that are the product of a local environment, or do they have more general currency? In particular, the validation exercise sought to check the applicability of the cases used to generate the Australian standards framework, the three frameworks and also the dimensions that constitute these frameworks.

THE CONTEXT

The international setting selected for this validation study was New Zealand. There are many similarities between the two countries. In addition, both have undergone parallel processes of decentralisation of the educational structures that suggest the tensions, problems and dilemmas facing principals in each country might be comparable.

The Australian Context

The Western Australia state education system, like its counterparts in all Australian states, has been in the throes of major reorganisation for more than a decade. Since the beginning of the twentieth century, its centralised state school system, like all education systems in Australia, provided education of comparable quality across large and economically diverse states with unique demographic and social needs. Like policy in each Australian state, policy announced in *Better Schools in Western Australia: A programme for improvement* (1987) promised a change from a highly regulated and centralised education system to one in which schools would have control over their goals as well as the resources and strategies for achieving them. School-based management was to replace the traditional reliance on the centre so that schools would become more flexible in coping with change, more responsive to meet local needs, and more accountable in the effective and efficient use of their resources.

This reorganisation of education delivery is one of three reform cycles aimed at shifting control away from the centralised state education structures (Angus, 1995). The first reform effort occurred in the 1970s when a Labor federal government set up the Australian Schools Commission to distribute funds to states for special purpose programs for disadvantaged students and to promote innovation. For the first time some schools had authority to initiate and implement their own programs, applying for funds, preparing budgets and engaging in school-level decision-making processes. The second reform cycle, during the

1980s, came from state-initiated policies, regardless of the political party in power. The policies were characterised by strategies for getting better value for education expenditure, a common language of efficiency, accountability, quality and excellence (Beare, 1991) and a belief that devolving problem-solving responsibility to the school level would improve the quality of education (Angus, 1995). The third reform cycle had its impetus in a coalition of industry, unions and employers. The driving force, the Australian Council of Trade Unions, urged managers and workers to share responsibility in their work organisation, with the goal of increasing efficiency, productivity and international competitiveness. Yet, in spite of three nation-wide reform attempts during the last quarter of a century, the highly centralised structures of education Australian states have remained resilient.

The reform agenda itself rests uncomfortably between two competing philosophies: the social democratic tradition and corporate managerialism. On the one hand, the social democratic tradition, which underpinned the first reform cycle of the 1970s, promotes public education as a tool for social reform and economic opportunity for the less advantaged (Dow, 1991). Social democratic reforms press for decentralised structures in which schools are autonomous and teachers as professionals are involved in making important decisions (Meadmore, Limerick, Thomas & Lucas, 1995). On the other hand, corporate managerialism is concerned about economic rationalism (Knight, Lingard & Porter, 1993). The focus is on centralising structures and aligning schools with corporate management through bureaucratic regulation and centrally prescribed parameters. Its goals are efficiency, effectiveness, impartiality, and accountability. As succeeding government policies reflect elements of both philosophies, it is not surprising that principals find demands made by systems and school communities both baffling and confusing (Wildy, 1998, 1999; Wildy & Wallace, 1997).

This is the Australian context in which the case-based standards framework for school principals was developed. The following section provides an overview of the historical background to the New Zealand reforms and some of the issues facing principals.

The New Zealand Context

New Zealand schools are undergoing similar changes to those in Australia. Since 1989 New Zealand schools have been subjected to parallel political and economic forces toward decentralization. *Tomorrow's Schools* (Lange, 1988) gave formal responsibility to schools to administer their budgets to maintain the school property, deliver an appropriate curriculum and meet legislative requirements. Elected Boards of Trustees were to be responsible for the selection of the principal, as well as the configuration and appointment of staff. The principal was to be accountable to the Board for the management of the school. Not only were principals, as Board members, to play a major role in staff appointment, they were also to assume responsibility for professional development of teachers. This responsibility included setting up performance management systems and assessing staff against teaching standards. In effect, schools were to become self-managing.

However, while the thrust toward school site management is common to both contexts, the New Zealand school system has embraced professional standards for teachers and principals. The principals' standards cover professional leadership, strategic management, staff management, financial management and

statutory reporting requirements. The principal is accountable to the Board of Trustees and faces a mandatory annual appraisal against the Ministry of Education's Performance Standards (Ministry of Education, 1998). The Board has the authority to make a supplementary salary payment contingent on its assessment of the principal's performance.

Furthermore, unlike their Australian counterparts, New Zealand principals face explicit and compelling pressures of consumer choice and competition. Behind the public policy of decentralisation lie the forces of a market-driven economy, whose tenets of consumer choice and competition, it is argued, are precursors of effectiveness and efficiency (Codd & Gordon, 1991). Principals are increasingly called on to balance educational imperatives, market forces, political hegemony and managerial complexity (Robertson, 1991, 1992). Many schools have taken the opportunity to be entrepreneurial within their schools and communities. Business partnerships and sponsorships are common in schools. Equally common is the trend toward marketing for international and national students. Increased autonomy over educational expenditure is an aspect of the reforms that has allowed schools increased flexibility and responsiveness to local needs. Because schools have been given financial authority, their ability to prioritise expenditure in the interests of local needs has led to the emergence of unique school identities.

However, the accumulation of research evidence indicates that such advantages come at a cost to schools and to principals. There is evidence of a widening gap between rich and poor schools (Gordon, 1994; Thrupp, 1996), although some principals of lower socioeconomic schools welcomed the autonomy to channel resources to meet needs in their schools (Strachan, 1999). There are signs that the continued intensive workload is taking its toll on principals' energy, making the principalship less attractive as a career path for teachers (Wylie, 1997). The principals of small schools, often less experienced than principals of large schools, are vulnerable to community pressure (Harold, 1995). In addition, under public pressure to portray a positive impression, many school leaders are feeling isolated and suffering loss of confidence (Robertson, 1998).

The process for appraising the principal against the performance standards is seen to be time-consuming, if conducted effectively (Fletcher & Kedian, 1998). There are also deep concerns among principals when they are subject to appraisal by non-professionals who largely make up their Boards of Trustees (Piggot-Irvine, 1999). Researchers such as Fletcher (1999) and Robertson and Allan (1999) question the nature and role of the standards used to assess the performance of principals. For example, there appears to be an over-emphasis on routine managerial activity and regulatory compliance, and an under-emphasis on the development of values that foster a culture for learning in the school. In the face of an expanded workload and responsibility for new and more varied tasks, it is not surprising that many principals have become reactionary and focused on management roles, or fled the field.

It appears that the New Zealand and Australian principals face increasingly complex and diverse demands. Both enjoy the benefits of decentralisation but also suffer the consequences of expanded work loads and heightened accountability. Both face the conflicting push and pull of the system that employs them and the communities they serve. Self-managing schools offer opportunities to shape student learning programs to meet local needs. The opportunity to determine their unique identity is more evident for New Zealand schools that, for more than a decade, have had authority to appoint staff and to manage their own budgets. The consumerist pressures of the market philosophy seem also to place principals

under increased public scrutiny. However, the commonality in the contexts suggest that the New Zealand setting would provide an appropriate site for exploring the process of developing a case-based principals' standards framework.

RESEARCH QUESTIONS

This study explored the extent to which the method of developing the Australian case-based standards framework for school principals' performance was applicable in a context which had undergone similar school system restructuring. Three questions are posed:

1. How valid for the New Zealand setting are the cases on which the Australian framework is based?
2. How applicable to the New Zealand setting are the three frameworks—duties, interpersonal skills, moral dispositions—that comprise the Australian standards framework?
3. How applicable to the New Zealand setting are the particular duties, interpersonal skills and moral dispositions that comprise the Australian standards framework?

The first question tests the face validity of the cases on which the Australian framework is based. The second question is answered by determining whether the essential qualities that underpin the Australian standards are also evident in the way performance is described by New Zealand principals. The third question is answered by determining whether the cases rated by New Zealand principals can be arrayed on continua using the dimensions of the Australian frameworks.

RESEARCH METHODS

The study was conducted in one region of New Zealand during 1999-2000. The site was selected because it encompassed a wide range of school types in a diversity of settings from large urban secondary schools to small isolated rural primary schools. There was also a rich mix of ethnic groups and some special interest schools. Principals had a range of length of experience, from newly appointed principals to those close to retirement. A group of principals in the region had participated in a professional partnerships program (Robertson, 1999).

There were three phases to this study: trialing the Australian cases; developing New Zealand cases; and rating the New Zealand cases and analysing the ratings.

Phase 1: Trialing the Australian Cases

The first phase, conducted during September 1999, involved trialing the Australian cases with New Zealand principals to ascertain the relevance of the cases. The procedure involved a two-hour professional development workshop, replicating the procedures used to generate the Australian framework. Participants were an invited group of 10 New Zealand principals representing a range of gender, age, length of experience, type of school and school location. Each participant received the same set of five cases selected from the 74 cases used to generate the Australian standards framework during 1996-1997. They depicted

principals at work in primary and secondary schools, facing tensions, dilemmas and problems concerning staff, students and community members. In each case, a different policy issue was addressed. Participants were asked to rate the quality of the principal's performance depicted in the case. There were three rating scales: a duty, an interpersonal skill and a moral disposition. Workshop participants spent approximately 15 minutes alone reading and rating the cases. Participants then compared their ratings with a colleague. Following this, participants engaged in a facilitated group discussion of the quality of performance depicted in each of the five cases.

What became evident during this workshop was that, although the issues underlying the cases were recognisable to participants, the surface detail of the cases was unfamiliar. For example, one case described the decision made by a principal when rain from a cyclone made unsafe the only trees in the school grounds. Whenever Australian principals rated this case, the ensuing discussion led to heated debate. Local readers understood the flat barren inland Australian landscape, the transitory nature of principals' appointments in small isolated schools and the tension between duty of care, local consultation and centrally prescribed policy. Without this implicit knowledge and understanding, New Zealand principals were unable to rate this case.

The Australian standards framework used cases because they incorporate the "circumstance and timing" (Louden & Wildy, 1999a) of principals' work. Cases capture what long lists of duties and dispositions do not capture. They portray the detail of the particular situation, its point in time and place, and the rich texture of the livid experience of the principal at work. In this surface detail are embedded the values and beliefs about principals' practice. On the basis of workshop in New Zealand in September 1999, it was decided that the Australian cases lacked the face validity to make them applicable to the New Zealand principals. This decision led to the second phase of the study, the development of New Zealand cases.

Phase 2: Development of New Zealand Cases

The second phase was conducted during February 2000 and involved developing new cases in the New Zealand site. The research team (the authors) had credibility and rapport with school principals and were well versed in the local school-level issues around restructuring. The researchers had extensive experience in unstructured interviewing techniques and developing cases from interview data.

The process replicated the case development phase of the Australian study. A sample of 14 principals was selected which included women and men, rural (teaching) and urban (non-teaching) principals, indigenous leaders, primary and secondary schools in different socio-economic contexts. Each principal was interviewed for approximately one hour in their school offices. Two researchers attended the interviews, one interviewing and taking notes while the other listened and also took notes. All interviews were tape recorded and the tape recording was used to check the hand written notes. The interviews were unstructured and began by seeking general information about the school, the community and the principal's experiences. The researchers had prepared open-ended questions similar to those prepared for the Australian study, focusing on critical incidents in which practical problems or dilemmas were tackled. As in the earlier study, the prepared questions were not used. Principals had incidents to recount and did not require prompting or directing from the researchers. Following each interview, the researchers analysed the interview data, looking for

tension or dilemmas. Cases were created from the interview data, incorporating enough local detail to provide authenticity and enough structure to indicate an underlying tension or dilemma. This is a creative act of connoisseurship (Wildy, 1998). First drafts were returned to principals for amendment and to ensure that the accounts did not contain a breach of confidentiality.

The 14 cases were allocated to dimensions of the interpersonal skills and moral dispositions frameworks. By inspection, it was clear that the duties on which the Australian framework was based were different from the list of duties used to categorise New Zealand principals' work. Furthermore, the Australian study had shown that raters distinguished the quality of principals' performance in terms of interpersonal skills and moral dispositions rather than duties. The duties (such as staff management, financial management) provided the problem or territory in which interpersonal skills and moral dispositions were played out. The same sets of interpersonal skills and moral dispositions were evident regardless of the particular duty being performed. As a result, the duties framework was not used for rating the New Zealand cases.

Two frameworks – interpersonal skills and moral dispositions – were trialed. Two dimensions of the interpersonal skills framework – listening, and collaborating – were selected. Four dimensions from the moral dispositions framework were used: patience and persistence; fairness and consistency; sensitivity and tact; and courage and decisiveness. Each case was allocated to one of the interpersonal skills and two of the moral dispositions. Like the earlier Australian process, the rating task consisted of two parts. In one part, raters were asked to describe, in their own words, the characteristics of each principal's performance depicted in the case. Three marked spaces were provided on the rating sheet for each case. In the other part, raters were asked to judge the quality of the principal's performance on one interpersonal skill and two moral dispositions using a four point scale (poor, adequate, good, very good) with an "unable to rate" option.

Phase 3: Rating of Cases and Analysis of Ratings

The third phase was a replication of the ratings and analysis of the Australian study, using the locally generated cases. Two strategies were used to obtain ratings: a workshop and a mail survey. These strategies involved 34 principals each of whom rated the 14 cases and generated 476 responses, representing a 100% response rate. Of these, one third were obtained during a three-hour workshop, and the rest from the mail survey. Responses were anonymous. Raters came from a purposive sample of principals in the same region from which the cases were generated, selected to represent proportions of gender, school type and school location similar to the proportions in the school principal population of that region.

The data were analysed using two different strategies. The first strategy was a content analysis of the descriptors provided by principals in response to the cases. The aim was to determine what principals in New Zealand counted as good principal performance. The qualities that principals identify indicate the values underpinning the standards framework, that is, the essence of quality in the principalship (Louden & Wildy, 1999b). The second strategy was the use of the Extended Logistic Model of Rasch (Andrich, 1988; Wright & Masters, 1982) and the RUMM computer program (Andrich, Sheridan & Luo, 1997). This statistical

procedure is used to analyse the psychometric properties of scales by providing an estimate of the position of each case on a continuum and confirming whether cases grouped at the top, middle and bottom of a continuum represent approximately equivalent levels of competence. The aim was to establish whether New Zealand principal raters used the same scales as their Australian counterparts in making judgements about performances depicted in cases. In other words, the purpose was to test the applicability of the dimensions of the Australian standards framework in an international context.

Comparisons were made between the New Zealand and Australian results of the two data analysis strategies. The aim of the comparison was to ascertain the extent of the cross-cultural application of the three frameworks that make up the Australian standards framework, and the dimensions within the sets of duties, interpersonal skills and moral dispositions.

FINDINGS

The findings of this study are presented in three parts. First, there is a presentation of the content analysis of descriptors applied to the principals' performance depicted in the 14 cases. These are then compared with the Australian findings, to determine whether the essential qualities that underpin the Australian standards are also evident in the way New Zealand principals characterise performance. Second, the statistical analyses are presented and these are compared with the Australian findings to determine whether the cases can be arrayed on continua using the dimensions of the Australian frameworks. Third, two New Zealand cases are presented to illustrate the application of the framework in the international setting.

Values Underpinning the Standards Framework

Of a possible maximum of approximately 1,400 performance descriptors (three descriptors for each of the 14 cases from each of the 34 principal raters), there were 1,367 responses. When the lists of descriptors for each case were collated, it was found that only 323 different words or phrases occurred among the 1,367 responses.

Of these 323 different descriptors, the most frequently used were two words: collaborative and visionary. These two descriptors, nearly 1% of all the different descriptors, accounted for 9% of all responses. Each was used at least 50 times. The second group of most frequently used descriptors consisted of three words: determined, lone ranger and reflective. Together with the top two descriptors, these five words make up 2% of all the different descriptors and account for 18% of all the responses. Each was used at least 40 times. A third group of most frequently used descriptors of the quality of principals' performance consists of eight words: autocratic, caring, strong, sensitive, egotistical, ethical, decisive, and professional. Each word was used at least 20 times. Together with the first two sets of words, these make up 4% of all descriptors and account for 35% of all responses. Performance descriptors used at least 10 times make up a fourth set. When this fourth set is included, a total of 33 descriptors (10% of all the different descriptors used) accounts for more than half (54%) of all the responses. The groups of performance descriptors are summarised in Table 1.

Table 1: Principal Raters' Descriptors, by Frequency

Descriptor	Number or principal raters using descriptors			
	At least 50	At least 40	At least 20	At least 10
Collaborative, visionary	x	x	x	x
Determined, lone ranger, reflective		x	x	x
Autocratic, caring, strong, sensitive, egotistical, ethical, decisive, professional			x	x
Inexperienced, action-oriented, innovative, tries too hard, realistic, empowering, hard-working, high expectations, manipulative, supportive, responsive, strategic, courageous, lacking vision, unconfident, controlling, creative, fair, patient, reactive				x
Percentage of all descriptors	1	2	4	10
Percentage of all responses	9	18	35	54

The simple listing of the frequencies with which descriptors were used in response to the 14 cases indicates that a relatively small number of qualities (33) were identified by the 34 principal raters. When these words are gathered into clusters of synonyms and antonyms, the concentration on a small number of ideas about principals' performance is increased. In Table 2, the 33 most frequently used words are clustered into descriptors with similar or opposite meanings.

Table 2: Principal Raters' Descriptors, Clustered by Frequency

Quality	Synonyms and antonyms	Number of responses	Percentage of top 33 responses	Percentage of all responses
Collaborative	Autocratic, lone ranger, controlling, empowering	195	27	14
Determined	Strong, decisive, courageous, action-oriented, unconfident, inexperienced	158	21	12
Caring	Reflective, sensitive, supportive, responsive, patient	141	19	10
Ethical	Professional, tries too hard, realistic, hard worker, high expectations, manipulative, strategic, fair	141	19	10
Visionary	Innovative, creative, reactive, lacks vision	101	14	8
Total		**736**	**100**	**54**

When the performance descriptors are clustered, only five different concepts were used by respondents when they were asked to describe the qualities of the principal's performance depicted in the cases. Across the group of principal raters in New Zealand who responded to the cases, the values concerning the quality of principals' performance are generally consistent. The five qualities, in order of frequency, were collaborative, determined, caring, ethical, and visionary.

The descriptors used by principal raters in New Zealand and Australia to describe the characteristics of performance of school principals are similar but they also show some important differences. The following discussion relates to findings from the earlier Australian study (Louden & Wildy, 1999a). In both studies it was found that respondents used a small number of descriptors with a high level of frequency. For example, when 34 New Zealand respondents used 33 different words at least 10 times they accounted for 54% of all the 1,367 responses. When nearly 1,000 Australian respondents used 24 different words at least 30 times, they accounted for 33% of all the 3,822 responses. In both cases, there appears to be a common set of values shared among the cohorts of principal raters. Furthermore, when these sets of most frequently used descriptors are gathered into clusters with similar or opposite meanings, the concentration of meanings is significantly increased. For example, the 33 New Zealand descriptors reduced to five qualities, and the 24 Australian descriptors reduced to six different qualities. Table 3 contains the two sets of qualities and their synonyms and antonyms, listed in frequency order.

Table 3: Australian and New Zealand Principal Raters' Descriptors

Australian		New Zealand	
Quality	Synonyms & antonyms	Quality	Synonyms & antonyms
Supportive	Understanding, sensitive, patient, caring, selfish, concerned	Collaborative	Autocratic, lone ranger, egotistical, controlling, empowering
Decisive	Indecisive, weak, reactive	Determined	Strong, decisive, courageous, action-oriented, unconfident, inexperienced
Fair	Thorough, consistent, realistic	Caring	Reflective, sensitive, supportive, responsive, patient
Innovative	Visionary, thoughtful, proactive	Ethical	Professional, tries too hard, realistic, hard worker, high expectations, manipulative, strategic, fair
Collaborative	Autocratic	Visionary	Innovative, creative, reactive, lacks vision
Flexible	Inflexible, open-minded		

Nine descriptors appear in both sets of qualities: collaborative, autocratic, decisive, visionary, reactive, supportive, sensitive, caring, patient. However, there are five points of contrast which taken together provide insights into the differences between the two principal cultures.

First, perhaps the most striking difference was the relative importance of the quality of fairness. Australian principal raters used this word more frequently than any other word, 109 times among 3,822 responses. However, the word fair appeared only 10 times among the 1,367 responses from New Zealand principal raters. Fairness in the Australian context implies balance and reasonableness. When clustered with words such as thorough, consistent and realistic, fairness takes on a pragmatic connotation. However, in the New Zealand setting, fairness is associated with ethical practice and its opposite is manipulative.

Second, while the word collaborative appears in both lists, its location and clustering give different emphases. In the Australian list of qualities, collaboration was placed fifth in order, and was clustered with only one word, *autocratic.* In the New Zealand list, collaborative was the most frequently appearing word (64 occurrences among the 1,367 responses). Although it is clustered with empowering, it is also clustered with lone ranger, autocratic, egotistical and controlling which together account for 121 responses. While New Zealand raters

used collaboration as a quality for judging principals' performance, it seems to be its opposite, working alone, being a lone ranger, that occurred more frequently in the cases to which they responded.

Third, a similar argument can be made about the relative meaning of another quality that appears in both sets. Decisiveness and its opposite, indecisiveness, appeared as the second most frequently occurring quality among Australian respondents. Principal raters in the Australian study were critical of performance that lacked decisiveness, calling it weak or reactive. Decisiveness was noted in its absence, in the Australian setting. However, determination rather than decisiveness appears to be the more frequently used concept among New Zealand principal raters. The cluster of words consists of determined, strong, decisive, courageous, action oriented. Its opposite is not indecisive but unconfident and inexperienced.

Fourth, closely linked to the use of the word determined is the role of the word visionary, a word that occurred with the second highest frequency (53 times among the 1,367 New Zealand responses). Among the lists were other expressions such as "gets what [he] wants", "a mover and shaker", "clear goals", and "clear vision". In this context the word assumes an active dimension, in contrast to the Australian usage which is coupled with the more passive notion of thoughtfulness.

And finally, while Australian principal raters noticed the extent to which principals' performance indicated flexibility, this quality is not as frequently mentioned by their New Zealand counterparts. Being flexible is not seen by New Zealand principal raters to be a quality that distinguishes principals' performance. However, the need for creativity is evident in both cultures. Where creativity is paired with flexibility in the Australian principal perspective, creativity is linked with innovation and vision in the New Zealand perspective.

In summary, while there is substantial commonality among the Australian and New Zealand qualities, there are differences of tone and emphasis. Reforms of New Zealand education have opened the way for leaders to be visionary in setting up clear goals that lead to creating an identity for their school. Principals, in particular, feel the pressure to be entrepreneurial and innovative. Compliance with centrally formulated rules and regulations is less pressing than the need to develop locally shaped policies and procedures. It is not surprising that determination is rated highly as a quality of principals' performance.

The importance of this discussion for the development of the standards framework is that it suggests the dimensions of the framework that would be applicable in the New Zealand context. The dimensions of the moral dispositions framework would need to differ from the Australian dimensions in important ways. The dimension *flexible and creative* of the Australian standards would be replaced with *visionary and innovative* for the New Zealand standards. The dimension *fair and consistent* of the Australian standards would be replaced with *ethical and fair* for the New Zealand standards. The dimension *courageous and decisive* of the Australian standards would be replaced with *determined and decisive* for the New Zealand standards. The two sets of moral dispositions are listed in Table 4.

Table 4: Australian and New Zealand Moral Dispositions: Dimensions

Australian	New Zealand
Fair and consistent	Ethical and fair
Courageous and decisive	Determined and decisive
Flexible and creative	Visionary and innovative
Sensitive and tactful	Sensitive and tactful
Patient and persistent	Patient and persistent

In summary, the method of developing standards for principals' performance generated in Australia appears to have currency when applied to cases relating to the local New Zealand context. The content analysis described in this section reveals qualities of principals' performance that are similar to those identified in the Australian context. However, the analysis also reveals important differences between what is valued by Australian and New Zealand principals.

Scales of performance

The second type of analysis involved the application of the Extended Logistic Model of Rasch (Andrich, 1988; Wright & Masters, 1982) and the RUMM computer program (Andrich, Sheridan & Luo, 1997). Although the method is the same as that used in the Australian study in its use of Item Response Theory, the more accessible computer program, RUMM, was chosen over the Quest program that was used earlier. The aim is to place all the cases (items) from the interpersonal dimensions on one scale (continuum), and all the items from the moral dispositions on one scale, provided the items fit a Rasch (1960/1980) measurement model. The items that fit the model will appear in an order from easy to difficult, which when applied to the rating terminology appear as ranging from "easy to agree that this case depicts high quality performance" (high performance), through to "difficult to agree that this case depicts high quality performance"(low performance).

The model combines the goal of item scaling with extended response categories for items that measure attitude that are applicable to this study. Item difficulties and person measures are placed on the same scale. The Rasch method produces scale-free person measures and sample-free item difficulties. In other words, the differences between pairs of person measures and pairs of item difficulties are expected to be sample independent (Waugh, 1999). The RUMM program (1997) parameterises an ordered threshold structure, corresponding to the ordered response categories of the items. Thresholds are boundaries between response categories related to the change in probability of responses occurring in the two categories separated by the threshold. The RUMM program re-parameterises the thresholds to create an ordered set of parameters: location or degree of difficulty, scale and skewness. Location refers to average difficulty of the item on the continuum. Scale refers to the average spread of the thresholds of an item on the continuum. Skewness refers to the degree of modality associated with the responses across the item categories. Having calculated these parameters for the items, the program substitutes the parameter estimates back into the model

and examines the difference between the expected values and the observed values. Two tests of fit are used: the item-trait interaction; and the item-person interaction. The first uses a chi-square test of fit to examine the consistency of item parameters across the person estimates for each item. The second, the test of item-person interaction, examines the response pattern of persons across items and items across persons. The fit statistics approximate a *t* distribution with a mean of zero and a standard deviation of one. Values below zero indicate the response pattern fits the model too closely and suggest the present of dependencies. Values that are above zero indicate a poor fit to the model and suggest the presence of "noise".

The 14 cases were each rated by the 34 New Zealand principals on three of the six dimensions dimensions as shown in Table 5.

Table 5: Cases by Interpersonal Skills and Moral Dispositions Dimensions

Case	Listening	Collaborating	Sensitivity & tact	Fairness & consistency	Courage & decisiveness	Patience & persistence
A 2nd job for a principal		*	*			*
No-one talks about the pain of leadership		*	*	*		
Questioning competency	*				*	*
Berlin Wall		*		*		*
Artful resourcing		*		*		*
Many culture, many practices	*		*		*	
A hard act to follow		*	*			*
Ten foot tall and bulletproof	*		*	*		
Enabling power	*			*	*	
Making the dreams happen	*		*		*	
Baptism of fire	*			*	*	
"The buck stops here"		*			*	*

The data analysis indicated that two cases, *No one talks about the pain of leadership* and *Many culture, many practices*, did not fit the model and should be discarded. Of the six possible scales, two have reasonable reliability, two have low reliability and the reliability of two is too low. An example of a scale with reasonable reliability is the sensitivity and tact scale.

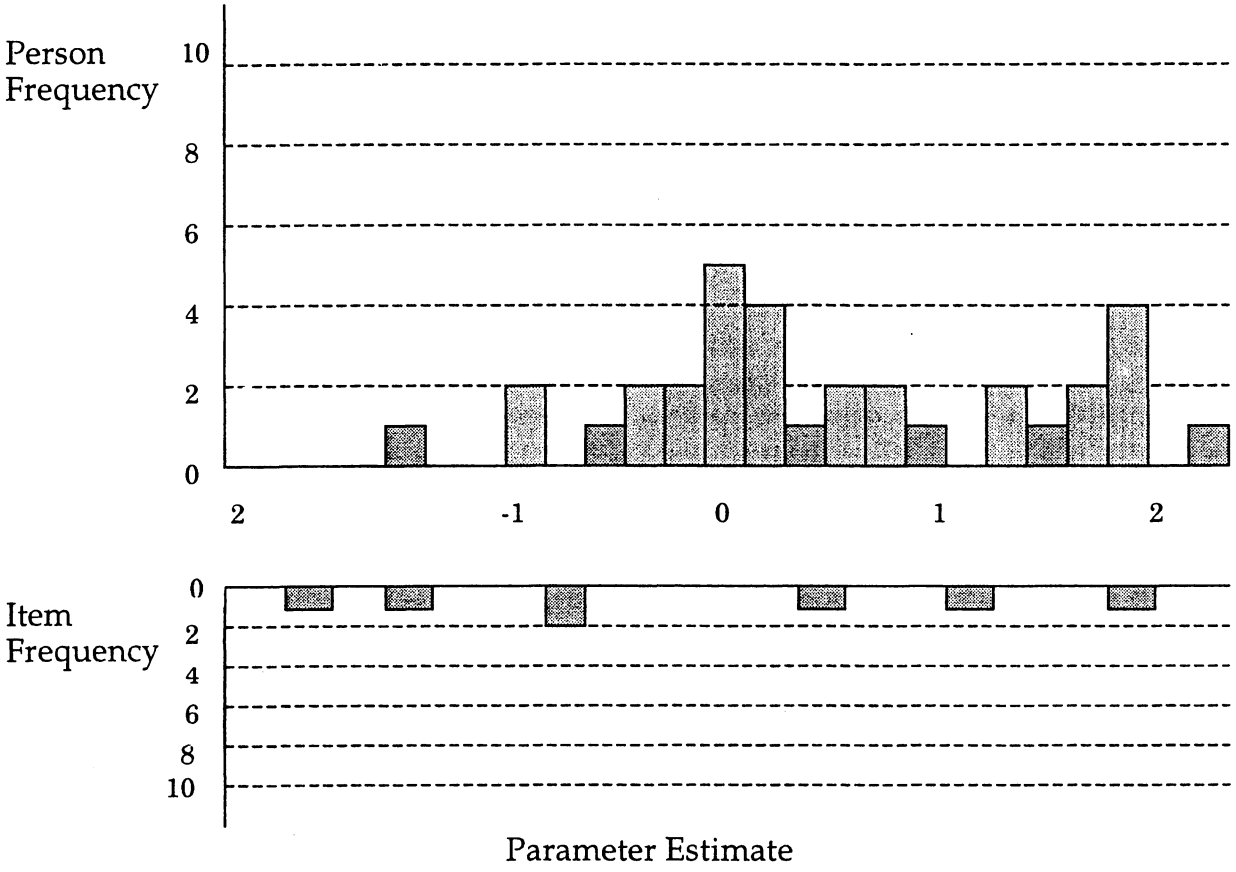

Figure 1: shows the graph of sensitivity and tact and the location of cases on the same scale.

From left to right on the continuum are these seven cases. The cases are listed in order of quality of performance from high performance to low, on the dimension, sensitivity and tact.

High	Many culture, many practices
	No one talks about the pain of leadership
	Making the dreams happen
	A hard act to follow
	A second job for a principal
	Ten foot tall and bulletproof
Low	"Step out of the way"

The items (cases) on this scale have a good fit to the measurement model, indicating strong agreement among the 34 raters about the location of the cases on the scale. The item-rater tests of fit (Total item Chi Sq 7.914; Pearson separation index 0.568; Test of power of fit Reasonable) indicate that there is good consistency of rater with underlying measurement scales and item response patterns. Table 6 shows the location of items on the continuum, the fit with the model and the probability of fit under the measurement model. Table 7 shows the Threshold values of the seven items fitting the model.

Table 6: Location of Items on the *Sensitivity and Tact* Continuum, with Indicators of Fit

Label	Location	SE	Residual	ChiSq	Probab
Ex004 I004	-0.234	0.26	0.164	0.012	0.994
Ex001 I001	0.502	0.21	0.738	0.727	0.687
Ex002 I002	-0.155	0.21	0.105	0.825	0.653
Ex003 I003	-1.780	0.32	-0.022	0.900	0.628
Ex005 I005	1.758	0.23	-0.102	0.905	0.626
Ex006 I006	1.134	0.21	0.788	1.241	0.525
Ex007 I007	-1.226	0.36	0.814	3.303	0.170

Table 7: Threshold Values of *Sensitivity and Tact* Item Response Categories

	1	2	3
Ex001 I001	-.196	-.013	-.256
Ex002 I002	-0.155	0.21	0.105
Ex003 I003	-3.673	-.793	-.873
Ex004 I004	-1.346	-.781	1.425
Ex005 I005	.693	1.202	3.379
Ex006 I006	.447	1.294	1.662
Ex007 I007	-4.523	-.561	1.407

The RUMM analysis of ratings of performance depicted in 14 New Zealand cases has been illustrated in relation to one of the six scales. The contribution that this kind of statistical analysis makes to the discussion of principals' performance standards concerns the issue of variation in performance on particular dimensions. As was found in the earlier Australian study, cases depicting principals' performance can be arrayed along continua to demonstrate levels of performance. What this New Zealand study demonstrates is that some, but not all, of the dimensions generate scales of performance. The dimension that fails to discriminate performance is fairness and consistency. It is likely that, as discussed in the previous section of this Findings section, consistency is not a value on which New Zealand principals rate performance. A more appropriate dimension for the New Zealand context would be *fair and ethical practice.* Further iterations of ratings and analyses would be needed to test this proposition.

APPLYING THE AUSTRALIAN FRAMEWORK TO NEW ZEALAND CASES

In the next section, two of the cases on the sensitivity and tact scale, *A hard act to follow* and *Ten foot tall and bulletproof,* are presented. The first case is an example of moderately high performance and the second case represents poor performance. Both are located on the same scale, sensitivity tact, a dimension of the moral dispositions framework that has currency in both Australian and New Zealand contexts. (Item labelled 1004 in Tables 6 and 7 is the case *A hard act to follow;* item labelled 1006 is the case *Ten foot tall and bulletproof.*)

The first of the two cases, *A hard act to follow,* is an example of a case that drew mixed evaluations on the three dimensions on which it was rated. On the interpersonal skills dimension of collaborating, this case was located in the middle of the scale. On the moral disposition sensitivity and tact, this case was located in the middle of the scale. However, on the moral disposition patience and persistence, *A hard act to follow* was located at the top, or high end, of the scale. The case *A hard act to follow* follows.

A hard act to follow

My first principalship was in a high performing school. It was a very old school in a high socioeconomic area. All its former principals had high local and national profiles and many years of experience. The Education Review Office had given the school a clean slate. I was appointed to maintain the school. However, I could see things had to change. After five years, I have made those changes but it was hard. I had to fight to survive. I fought to build my credibility and to improve the school.

I found that staff were accustomed to being told what to do. My predecessor implied: "You don't have to like it. Just do it." Most of the work was done in the principal's office. This was not my way. I am a team person not a front runner. I want to empower staff. Besides, the former principal worked till midnight every night. I knew I couldn't do everything nor was I prepared to spend that amount of time at work.

I recall my first staff meeting. I posed questions, thinking we would engage in a professional debate. There was silence. No one was prepared to talk. It was a very short meeting. Later I heard there was lots of talk among staff and parents, all criticising me for not giving directions. They said to me: "You are the principal. Your job is to tell us what to do." The car park was full of talk. There was a group who wanted me to make dramatic changes but I didn't change a thing in the first six months.

Gradually I revved up the teachers. I started with the curriculum. I developed systems of support such as curriculum teams with their own budget and decision-making authority. Each team has now collaborated to produce a program for each curriculum area, a seamless document, up and down the school. The Assistant Principal and the Deputy Principal each has responsibility for half the curriculum and I take responsibility for those specific areas targetted for review. Every teacher belongs to a team, either as a leader or a member. I am in the Information and Communication Technology (ICT) team and take part in all their professional development activities because I need a better

understanding of ICT. I am showing that we are all leaders and learners.

Class allocation was always done by the former principal during the summer break. Teachers and students arrived at school and in assembly on the first day class groups and their teachers' names were read out. I wanted students and teachers to know the groupings before the end of the school year. I started working with individual teachers about their preferences and their recommendations based on students' performance and their social needs. This was a complicated job as enrolments were always growing. But I was adamant that all students would go home knowing the name of their teacher for the following year. Unexpectedly, on the last few days of the year, another 20 students were enrolled, disrupting all our previous groupings. I pressed the Assistant Principal to redo the classes. He was upset. I insisted it was done before the last assembly. In times like this I think: 'Maybe it is better just to do it myself.' It was also confusing for him when I assumed this autocratic stance, after we had been working collaboratively during the year.

Last year there was another staff meeting in silence. I was stunned. I wondered if I was back to square one. Then I realised what was happening. The teachers were exhausted. Talk always seemed to lead to new ideas and that meant more work. Now I take account of teachers' well-being and help them to monitor their own commitments. I am learning to develop professional support networks to help me monitor my own well-being.

This case, *A hard act to follow,* shows a rich mix of each of many qualities: collaboration, care, professionalism, determination and vision. More than a quarter (28%) of the 96 descriptors applied by the 34 principal raters who responded to this case related to the principal's collaboration. For example, respondents noted that the principal was team oriented, inclusive, consultative, empowering and co-operative. Only one respondent gave an opposite view describing the performance as top down. One fifth of the respondents described the performance as caring, sensitive, gentle, patient and reflective. Another fifth of respondents described this performance in terms of its professionalism using terms such as hard working, realistic, consistent. However, when respondents noted that this principal was visionary they also commented that the principal was confused and not a good change manager. Similarly, while some noticed the principal's determination and decisiveness, others drew attention to the principal's feelings of insecurity.

What appears to be valued by principal raters is the principal's commitment to make changes that were not only in students' interests but also in the interests of teachers. Clearly, this principal had a vision for the school that was at odds with the judgement of the Education Review Office. Tackling the complacency of the staff and parents would take courage and also patience and sensitivity. When principal raters rated this case moderately high on collaborating, they may have been drawing attention to the apparent inconsistency when, under pressure, the principal adopted an uncharacteristically autocratic stance. Perhaps when the principals rated the performance only moderately high on sensitivity and tact they were drawing attention to the principal's initial failure to recognise teachers' exhaustion. Similarly, when principals rated the performance high on patience

and persistence, they may have been impressed by the principal's commitment to change in the long term, even at the expense of personal discomfort.

The case, *A hard act to follow,* depicts all of the qualities that New Zealand principal raters consider essential in distinguishing the performance of principals. In addition, the performance of this principal illustrates the balance between competing qualities: between being collaborative and having a vision; between caring for others and taking action; and between patience and determination. This is an example of a case that has been rated using the dimensions of the Australian standards framework.

The second of the two New Zealand cases, *Ten foot tall and bulletproof,* has been rated using three different dimensions of the Australian standards framework. The case was rated on the interpersonal skills dimension of listening, and on two moral disposition dimensions: sensitivity and tact, and fairness and consistency. On all three scales, this case was positioned among the lowest two items. The case *Ten foot tall and bulletproof* follows.

Ten foot tall and bulletproof

I was tapped on the shoulder and asked to apply for the principalship of a school that was going through a rough time. Its name had been dragged through the mud and enrolments were falling. There were questions of the suitability of its female principal and the Board of Trustees. I was singled out because I had been successful earlier in my career in turning around a school in a similar situation. I took on the job of teaching principal in this middle class rural school. I began by finding out the failings of my predecessor. It seemed that staff, parents and students were looking for strong leadership. I aimed to build staff morale and to develop students' self esteem. But first I had to establish myself as the leader.

I use a non-confrontational approach but I am hard underneath. For example, a teacher complained to me that a member of the Board of Trustees was making comments about the standards of teachers' dress. He singled out two female teachers for particular criticism. The Board member would be in the staff room at the end of the school day criticising teachers for not being in their classrooms. I spoke to the Board member and told him not to talk to the teachers in that way. Then I spoke to the staff. I said: "If he comes to the staffroom and criticises teachers' dress or behaviour, don't react. Just look him in the eye and then walk out of the room." This is what they did. After a time, his comments reduced. He has since left the Board.

I have continued to have my disagreements with the Board. For example, I don't sit comfortably with the Board judging the performance of principals. They don't have the knowledge or the skills especially when they are non-professionals or unemployed people. How can they make decisions about the salary and professional development needs of the principal when they do not even have a job themselves? I am the professional leader of the school. You need a Chair who knows this and will listen to you.

There are times when you have to employ tactics that are not quite kosher. For example, I identify who I want on the Board and approach them and ask them to apply. I work out who would make a good Chair.

> It must be someone with whom I can build a good relationship. When I first came, one parent came into my room and told me she was the community stirrer. I turned and looked at her and said: "Now I'll know where all the flak is coming from." I knew she was a gossip and soon she would spread the word that I was a principal who would give as good as he got. I know how to use the community as a vehicle to get what I want. But you have to be ten foot tall and bulletproof. I have high standards and I expect them to be met. If teachers don't shape up, I call them in and have a word. I don't judge people too often but when I do I choose my words carefully. We have a stringent performance management structure involving peer appraisal and professional partners. I developed this process collaboratively using subtle manipulation so that I got what I wanted. It is working well. We haven't had any competency issues.
>
> In the past three years since my appointment, numbers have grown from 130 to over 400. This is a dramatic increase. We have new buildings and extra staff to accommodate the growth. It is an expanding area and there have been boundary changes. Both of these factors have contributed to our growth. But I think my strategies have paid off too.

The performance depicted in *Ten foot tall and bulletproof* illustrates the tension between three of the essential qualities: collaboration (in this case, its opposite, autocratic practice), determination, and ethical behaviour. Of the descriptors selected by principal raters to apply to this performance, almost one half of the 100 words could be grouped as either autocratic or determined. For example, 24% of the descriptors were synonyms of autocratic such as controlling, loner, and arrogant. Another 21% of the descriptors were synonyms of determined such as strong, decisive and directing. There was another cluster of descriptors that were synonyms or antonyms of ethical such as manipulative, pragmatic, high standards. This cluster consisted of 21% of the descriptors. The importance of this set of words is that it suggests that principals who rated this case are concerned about the ethical dimension of the performance depicted in the case. Here, more than in the Australian cases, when raters are asked to make judgements about fairness, they are less inclined to comment on principals' performance in terms of being fair and consistent as being fair and ethical.

The tension seems to lie between the goal of turning around a failing school, and the means by which the principal achieves this goal. On the one hand, the principal is appointed on the basis of previous success in turning around a failing school. By the principal's account, the methods are again successful, although it might be argued that the increased enrolments are the outcome of school boundary changes in an expanding community, rather than the direct consequences of the principal's strategies to increase enrolments. On the other hand, by the principal's own admission, the methods used are tactics that are "not quite kosher" and include "subtle manipulation". These strategies include identifying a suitable community member to be elected as the Board chair, confronting parents, and using the local network to seed ideas about the new principal's style and expectations. The principal's key strategy, it seems, is to provide strong leadership to get what he wants for the school, even if this entails being "ten foot tall and bulletproof". Indeed, before building staff morale and developing students' self esteem, this principal first aims to assert himself as the leader.

In projecting an image of "ten foot tall and bulletproof" the principal shows that he is "hard underneath", willing to take on a Board member's criticism of teachers' behaviour and dress, and disagreeing with the Board's formal responsibility to judge the principal's performance. Furthermore, the principal sets stringent performance standards for staff and has managed to get teachers to shape up without having to take up competency procedures.

However, principal raters consistently rated this performance low. The principal's performance is low on dimensions of listening, sensitivity and tact, and fairness and consistency. The New Zealand principals who rated this case suggest that achieving the goal of turning around a failing school is not justified at any price. Being a strong leader with determination to achieve a clear goal at the expense of collaboration, care and ethical practice, it appears, is not acceptable in this New Zealand principal culture.

This second case, *Ten foot tall and bulletproof*, illustrates the application of the Australian standards dimensions of listening (one of the interpersonal skills) and sensitivity and tact (one of the moral dispositions). What is more interesting is that the case also illustrates the subtle but important distinction between the two principal cultures in their understanding of the concept of fairness in their daily work.

CONCLUSIONS AND IMPLICATIONS

This article has presented an application in an international setting of an innovative method of using cases, short narrative accounts, to develop a standards framework for school principals' performance. The application was small compared with the earlier Australian study. The New Zealand study, conducted during a six month period, involved 14 cases generated in one school region, rated by 34 principals on six dimensions of performance. The Australian study, conducted over two years, involved 74 cases generated in three regions, rated by 1,000 school administrators on 15 dimensions of performance. Despite the relative small size of the study, its findings are quite clear. The case-based approach to developing standards is applicable in a setting that has some similar features but many distinct differences.

This study supports three conclusions. First, cases generated in one context are not applicable in a different cultural setting but that the method of developing cases is readily applied cross-culturally. Second, the values underpinning the framework developed in Australia are similar to, but not the same as, those about which principals in New Zealand assess principal performance. However, the cases can be used in this way to identify the values that underpin the work of principals. Third, there are some similarities in the particular dimensions on which the framework is grounded. However, the particular dimensions can be identified, in an iterative way, through analysis of the qualities that underpin the work of principals. What this study has shown is that the method of developing standards is robust. What is also evident is that, because the method is built on cases and cases capture circumstance and timing, that is the local and particular in principals' work, the materials on which the standards are based need to be developed locally. Most importantly, this study demonstrates that there are enduring tensions and dilemmas that challenge principals, regardless of the context in which they work.

The experience of working across two cultures revealed that principals who were involved in the study reported that the experience was professionally

enriching. The researchers noted the enthusiasm with which principals agreed to be interviewed and their reluctance to close the conversation at the end of the planned one-hour appointment. Principals were keen to participate in the rating workshop. All of the principals invited to take part in the mail survey strategy completed the task and returned the 14 cases. The commitment to this research is more noteworthy in the face of principals' unquestionably heavy workload. It is likely that the research activity provided principals with a respite from their intense, action-oriented day. Further, it is possible that the opportunity to talk about their daily experiences and to read another person's interpretation of their practice provided some kind of validation of their work. These are speculations. What was clear, though, was that principals who participated in the research were learning about their own practice. For example, one principal wrote on the bottom of the last of the rating sheets, "I wish I had seen some of these five years ago. It would have changed how I approached some things in my work". The New Zealand study has confirmed what was revealed in the Australian research, and reinforced on countless subsequent occasions. There is no doubt that reflection on cases depicting principals grappling with the dilemmas and problems of their everyday work is professionally rewarding for principals.

This study has also demonstrated the importance of attending to interpersonal skills and moral dispositions. In increasingly decentralised school structures, principals are thrust more closely into contact with their stakeholders, particularly parents and community members. New Zealand principals need the skills and understanding to work effectively with their Boards of Trustees, particularly with the chair of the Board. Similarly, as more and more decisions are devolved to the school site, principals face increasingly the need to recognise, understand and balance the tensions of responding to competing, but equally plausible, demands. For example, the tension between competition and collaboration was never more evident than in workshop discussions of the issue of falling roles. The tension between building up enrolments, on the one hand, and closing small specialist schools with falling enrolments, on the other hand, became apparent when the principals of such schools sat in one room and debated issues as they appeared in the cases. It was clear that principals valued the determination to succeed of those whose vision was to turn around a school with declining enrolments. However, when principals met colleagues whose schools were being closed because of declining enrolments and listened to their commitment and dedication, the issue seemed less clear-cut. In their efforts to handle and resolve such tensions, principals need highly developed interpersonal skills such as listening, collaborating, negotiating and the ability to confront the complexity of the situations they face. They also need to balance competing pressures towards sensitivity and patience, on the one hand, and determination and vision, on the other hand. More important, though, is the skill and wisdom to arrive at a point of resolution that is, above all, ethical. Standards frameworks that ignore these core skills and dispositions fail to capture the heart of principals' work.

Many school systems are moving towards standards-based approaches to educational improvement. Too often, the attempt to provide universal standards eliminates the contextual detail that is essential in describing and judging performance. By using cases rather than lists of duties to describe the standards, the framework developed in Australia and validated in New Zealand captures standards in the context of everyday work. International testing of the cases and continua was an important step in determining the robustness of the method of generating case-based standards.

The next step in the research programme is to develop authentic assessment strategies to allow others to make judgements about performance using the case-based standards framework. This standards framework identifies the qualities of good performance and also the way performance varies on some dimensions. However, the framework is not an assessment instrument. New research will focus on assessment strategies such as portfolios and simulation activities—currently used to measure principals' performance on practical duties—and the extent to which these assessment strategies are sufficiently reliable and valid to allow others to make judgements about performances in terms of interpersonal skills and moral dispositions. Building on work done by the National Association of Secondary School Principals (NASSP, 1997) and the Council of Chief State School Officers (CCSSO, 1996a) in setting and assessing professional standards for school principals, the next phase of this study will develop assessment procedures that link to a case-based standards framework involving interpersonal skills and moral dispositions.

REFERENCES

Andrich, D. (1988). A general form of Rasch's Extended Logistic Model for partial credit scoring. *Applied Measurement in Education, 1*(4), 363-378.

Andrich, D., Sheridan, B., & Luo, G. (1997). *RUMM: A windows-based item analysis program employing Rasch Unidimensional Measurement Models.* Perth: Murdoch University, School of Education.

Angus, M. (1995). Devolution of school governance in an Australian state school system: Third time lucky? In, D. Carter & M. O'Neill (Eds.), *Case studies in educational change: An international perspective* (pp. 6-27). London: The Falmer Press.

Beare, H. (1991). The restructuring of schools and school systems: A comparative perspective. In, G. Harman, H. Beare & G. F. Berkeley (Eds.), *Restructuring school management: Recent administrative reorganisation of public school governance in Australia* (pp. 14-25). Canberra, Australia: The Australian College of Education.

Codd, J., & Gordon, L. (1991). School charters: The contractualist state and education policy. *New Zealand Journal of Educational Studies, 26*(1), 21-34.

Council of Chief State School Officers. (1996a). *Interstate School Leaders Licensure Consortium. Standards for school leaders.* Washington DC: Author.

Council of Chief State School Officers. (1996b). *Interstate School Leaders Licensure Consortium. Standards for school leaders. Draft model standards,* Washington DC: Author, p. 2.

Dow, G. (1991). *What do we know about social democracy?* Paper presented to National Social Policy Conference, Social Policy Centre, University of New South Wales, Australia.

Fletcher, M. (1999). *A Headteacher's lot! Reviewing a principal's performance-what should the focus be on?* Paper presented at the BEMAS Conference, Manchester, September.

Fletcher, M., & Kedian, J (1998). *Appraisal: Towards becoming an effective learning organisation.* Paper presented to NZCTE Conference. Hamilton. October.

Gordon, L (1994). Rich and poor schools in Aotearoa-New Zealand. *New Zealand Journal of Education Studies, 29*(2), 113-125.

Harold, B. (1995). *Negotiation, contestation and conflict: The introduction of new educational policies to three rural schools.* Unpublished doctoral dissertation: University of Waikato, Hamilton.

Knight, J., Lingard, B., & Porter, P. (1993). Restructuring schooling towards the 1990s. In B. Lingard, J. Knight & P. Porter (Eds.), *Schooling reform in hard times* (pp. 2-22). London: The Falmer Press.

Lange, (The Hon.) D. (1988). *Tomorrow's schools: The reform of educational administration in New Zealand.* Wellington: Government Printer.

Louden, W., (in press/2000). Standards for standards: The development of Australian professional standards for teaching. *Australian Journal of Education, 44*(2).

Louden, W., & Wildy, H. (1999a). 'Circumstance and proper timing': Context and the construction of a standards framework for school principals' performance. *Educational Administration Quarterly, 35*(3), 399-422.

Louden, W. & Wildy, H. (1999b). Short shrift to long lists: An alternative approach to the development of performance standards for school principals. *Journal of Educational Administration, 37*(2), 99-120.

Meadmore, D., Limerick, B., Thomas, P., & Lucas, H. (1995). Devolving practices: Managing the managers. *Journal of Education Policy, 10*(4), 399-411.

Ministry of Education. (1987). *Better Schools in Western Australia: A programme for improvement.* Perth, Australia: Author

Ministry of Education. (1998). *Principal performance management.* Wellington, New Zealand: Author

Queensland, Department of Education. (1996). *The professional development framework for principalship.* Brisbane, Australia: Author.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests* (expanded edition). Chicago: University of Chicago Press.

Robertson, J. M. (1991). Dilemmas faced by school principals. *The New Zealand Principal, 6*(3), 17-18.

Robertson, J. M. (1992). Qualities of principals. *The New Zealand Principal, 7*(2), 13-14.

Robertson, J. M. (1998). From managing impression to leadership perspectives. *International Journal of Educational Research, 29*(4), 359-370.

Robertson, J. M. (1999). *Principals working with principals: Keeping education at the centre of practice.* SET, Number One, Item 9. Wellington: New Zealand Council for Educational Research.

Robertson, J. M. & Allan, R. (1999). *Teachers working in isolation? Creating opportunities for professional conversations.* SET, Number Two, Item 3. Wellington: New Zealand Council for Educational Research.

Piggott-Irvine, E. (1999). Making principal appraisal work. *The New Zealand Principal, 14*(3), 10-12.

Strachan, J. (1999). Feminist educational leadership in a neo-liberal New Zealand context. *Journal of Educational Administration, 37*(2), 121-138.

Teacher Training Agency. (1996). *Key principles and draft national standards for new headteachers.* Chelmsford, Essex: Author.

Thrupp, M. (1996). *The school mix effect.* Unpublished doctoral dissertation. Victoria University of Wellington, Wellington.

Waugh, R. (1999). Teacher receptivity to a system-wide change in a centralised education system: A Rasch Measurement Model analysis. *Journal of Outcome Measurement, 3*(1), 71-88.

Wildy, H. (1998). School principals and the dilemmas of restructuring. Unpublished doctoral thesis. The University of Western Australia.

Wildy, H. (1999). Restructuring and principals' power: Neither *freedom from* nor *freedom to. Leading & Managing, 5*(2), 114-124.

Wildy, H., & Louden, W. (2000). School restructuring and the dilemmas of principals' work. *Educational Management and Administration, 28*(3), 173-184.

Wildy, H., & Wallace, J. (1997). Devolving power in schools: Resolving the dilemma of strong and shared leadership. *Leading and Managing, 3*(2), 132-146.

Wylie, C. (1997). *Self-managing schools seven years on: What have we learnt?* Wellington: New Zealand Council for Education Research.

Wright, B. D., & Masters, G. (1982). *Rating scale analysis.* Chicago, IL: MESA Press.