Chapter 3: Levelled curricula, learning progress and skills tests

> In studying the child the teacher tries to learn his innermost thoughts so that she may be able to render her guidance intelligible to him. As she learns to understand him she begins to sympathize with him, and in return she secures his love; once his love is secured, he will follow her to the end of the earth, and the examinations will take care of themselves. Thus the weight of oppression becomes removed from the child; he becomes free and happy in his freedom, and the school is converted into the loveliest of homes.
>
> Rice, 1893, p. 97.

The previous chapter established that assessment schemes using scales and examples of known scale quality to estimate the learning status of students had their roots in the work of mainly American educationalists in the early 20th century. This chapter addresses the development of the Australian curriculum and assessments structures in the 1980s and 1990s, particularly as they impacted South Australia.

One distinctive element of the assessment process for these curricula was the requirement for teachers to judge how students were progressing. They were required to make these judgements using a curriculum framework structured in levels. This structure was a change from most previous curriculum descriptions, which described content to be covered in specific grades or Year levels. The level concept, a description of skills and other attributes classified into levels of increasing complexity, was key to the creation of the teacher judgement data, summarised in Chapter 7. In the mid 1990s South Australia also introduced statewide testing in primary schools. Data from the tests of the period feature in Chapter 6, and a comparison of teacher judgements using levels with the test assessments is made in Chapter 8.

The chapter outlines a brief history of the development of the levels structure for the proposed but only partly implemented Australian national curriculum of the mid 1990s and the strong link in that curriculum approach to teacher judgement as one basis for assessing student progress. South Australia contributed to the national developments while establishing its own level structure in the form of attainment levels, prior to the ultimate adoption of the national level structure.

The development of 'Profiles' and 'Levels' for Australia

Documents describing a level approach to curriculum were developed in the period from 1988 to 1993 under the auspices of the Australian Education Council (AEC) based on

recommendations from the Directors General Conference and the Australasian Cooperative Assessment Program (ACAP) (Lokan, 1997).

Progress for a student through the curriculum was described as eight levels of increasingly sophisticated skills and knowledge within eight learning areas. For each learning area there were two documents produced; a statement that provided "a framework for curriculum development" and a profile that "described the progression of learning typically achieved by students" (Curriculum Corporation, 1994c, p. 1)[10].

Aspects of the Statements and Profiles for Australian Schools (SPFAS) documents development process have been recorded by Lokan (1997), Marsh (1994), Piper (1997) and more informally by Jenkin (1996). Jenkin (1996) was well positioned to observe the developments in his role as executive officer to some of the committees involved in the developments nationally and in South Australia. Boomer, as Chair of ACAP and South Australian Associate Director General - Curriculum and formerly Director of the national Curriculum Development Centre played a key role in the initial development of the statements and the concept of teacher judgement assessment as an alternative to system wide and national testing (Jenkin, 1996). Boomer helped link local SA initiatives in attainment levels to the national approach.

Precursor approach- Attainment Levels in South Australia

During 1990 South Australia had been considering its own approach to a levelled curriculum in the form of attainment levels. Boomer commissioned the South Australian Curriculum Directorate to develop attainment levels (Education Department of South Australia, 1992) with the endorsement of the then SA Director General, Boston[11] (Jenkin, 1996). The brief for developers required 6 levels from Reception (R) to Year 10. Levels attained were to be standards referenced. Reports for parents based on these levels were to be provided along with a school and system perspective by curriculum area (Stehn, 1997).

Masters (1999) reports that Griffin's Literacy Profiles (Griffin, 1990) were "influential in shaping later initiatives to develop the South Australian Levels of Attainment and more

---

[10] For convenience these two aspects for each learning area are described generally as Statements and Profiles for Australian Schools (SPFAS) throughout the text, as introduced in Chapter 1.

[11] Boston subsequently moved to NSW as Director General. As chair of the AEC Curriculum and Assessment Committee (CURASS) he brought the national SPFAS document development program to its completion (Marsh, 1994).

significantly the national profiles" (p. 289). Both Masters and Griffin were influenced in their views on curriculum, learning and assessment by their understanding of the Rasch model (Masters, 2005; Griffin, 1998), underlining the conceptual impact of the Rasch model on aspects of the design of the South Australian attainment levels and ultimately the national profiles.

The implementation of attainment levels in 1992 meant that SA teachers were already exposed to some developmental concepts in curriculum design that would flow, in general terms, into the levels of the national profiles (Jenkin, 1996). The spin-off from the attainment levels to classroom assessment was less well developed. However, prior exposure to a similar concept meant that the lead-time into the eventual adoption of the profiles and levels approach in South Australia had a benefit of three years more exposure than applied for some other parts of Australia (Marsh, 1994).

The parallels in the two approaches are reinforced in a pamphlet of the SA Curriculum Division (Education Department of South Australia, 1993). The similarities justified the continuing "familiarisation programs for the attainment levels" in 1993 "when we are moving towards the adoption of national profiles". The similarities were that both were standards referenced, both valued teacher judgement, both provided a tool for teachers to describe student achievement and both were a description of the progression of learning typically undertaken by students in each learning area. Many SA teachers had a benefit in developing their understanding of national profiles from their experience with attainment levels. Industrial action relating to teacher work load, described briefly later in the chapter (Stehn, 1997), influenced which teachers were able to maximise their use of attainment levels. At some sites attainment levels and then SPFAS were effectively banned.

Teacher judgement in Attainment Levels and SPFAS

Boomer believed in a system of assessment and reporting that supported the use of teachers' judgments and that valued their tacit understandings of their students. He had in mind, according to Jenkin,

> a system that would help teachers make more reliable judgments about students' achievements … without undermining their professional credibility and integrity. Accordingly the profiles framework was seen as needing to accommodate the way teachers and students actually worked together in schools and to be sufficiently broad not to impose a construct that limited the classroom options. (Jenkin, 1996, no page reference)

ACAP with Boomer as chair, proposed standards referenced frameworks based on the work of Sadler (1987). While generally similar to criterion referenced reporting, standards referenced frameworks were seen as drawing "upon the professional ability of competent

teachers to make sound qualitative judgments of the kind they make constantly in teaching" (Sadler, 1987, p. 193). This capacity for qualitative judgment was seen as being able to be "refined to the point where it could be used, directly, in the classification of student achievement into grade levels" (Sadler, 1986, quoted in Sadler, 1987, p. 193). Implicit in the Sadler concept was a holistic judgement of the quality of student leaning, not based on tests or mechanically accumulated ticks of outcomes met.

The adequacy of the level descriptions

The correct linking of outcomes to levels was required for the SPFAS to be effective. The Australian Council for Educational Research conducted calibration studies in 1992 and 1993. The studies considered the separation of levels, the equivalence of outcomes within a strand within a level and the ability of teachers to understand the assessment process (Lokan & Wu, 1997). The studies showed consistency in the pattern of the upper and lower thresholds[12] of outcomes by profile level and an even growth from one level to the next.

Masters, while acknowledging that the national curriculum was "developed hurriedly, and was best viewed as drafts of frameworks for curriculum, assessment and reporting" (1999, p. 280), suggested that the materials provided opportunities to address a number of empirical issues. In particular he asked, "Was the sequencing of outcomes along each strand, based on the experience of curriculum designers, consistent with the empirical ordering of assessment tasks designed to address those outcomes?" He saw an iterative process linked to parallel collections of test data as "useful for revising and refining the outcome frameworks" (Masters, 1999, p. 280). This never materialised.

The development of the statements and profiles was completed in June 1993 at a time of change of governing party in many state governments, which ultimately resulted in uneven implementation across the states. However, the initial development of the national (and Australasian) approach to a levelled curriculum was originally a cooperative initiative of the Directors General of each of the education systems of Australia, through their senior curriculum leaders, rather than an imposition from any outside entities.

Implementation in South Australia

The July 1993 AEC meeting agreed that any future publication of material would be the prerogative of each state and territory. This resolution was meant to imply neither

[12] Thresholds are at the point where students on average, when assessed for a specific outcome in a strand, move from 'hardly ever' to 'sometimes' category in that strand. The upper threshold is the point where students move from 'sometimes' to 'almost always'.

endorsement nor rejection, by the states and territories. South Australia proceeded to purchase copies of all documents for its schools. The professional development program for South Australian teachers had continued, flowing easily from the attainment levels approach to the SPFAS approach, although the transition was not without tensions. Marsh (1994) records that South Australia "made considerable advances with an implementation program, largely due to the three years' experience gained from developing and trialling state attainment levels" (p. 169).

Stehn (1997) comments, "perhaps the most difficult years in the history of the South Australian curriculum review and reform initiatives were 1992 and 1993. These were the years of metamorphosis … from attainment levels to nationally developed Curriculum Statements and Profiles" (Stehn, 1997, p. 173). The teachers' union had difficulty with perceived economic rationalist motives to the reform and as well the additional demands on teachers' time and a period of industrial unrest continued. Some schools struck deals to implement the profiles with staff, who were in many cases intellectually and pedagogically sympathetic with the general approach, seeing it as a re-packaging of current practice (Stehn, 1997). The union banned involvement in implementing either the attainment levels or the SPFAS.

Appreciating the complexity of the industrial situation it is not surprising that most SA departmental effort went into the explication of the impact on curriculum planning and attempting to ameliorate the perception of increased workload. As a result there was very little interest in exploring the broader student assessment possibilities including methods of recording a student's level status with greater refinement (author direct experience, 1994 to 1996).

Progress indications within a level

The intention to collect data on levels achieved by students in South Australia had been signalled early (Education Department of South Australia, 1993). However, progress within levels does not appear to have been discussed in operational detail by the South Australian implementation planners. This was revealed, by implication, in the arrangements that were eventually put in place for the collection of level data from teachers.

Progress within levels was an issue both for system-wide data collection and for teachers themselves. The time between attaining the criteria for one level and attaining the criteria for the next level was about two years for any student. How part progress towards the next level should be described, or if it was even desirable to do so, was left unspecified. Teachers, however, were adamant that some form of part progress record would be needed or else any data collections would misrepresent the learning that was occurring (Private conversations by

teachers with the author and negotiations with the teachers union in which the author participated, 1997).

The Victorians Curriculum Standards Framework (CSF), essentially equivalent to the SPFAS, addressed progress within a level by dividing the distance into three zones; 'beginning', 'consolidated' and 'established' (Department of Education, Employment and Training, Victoria, 1996). The South Australian Curriculum Division of the time was not enthusiastic about clarifying the detail of what the likely data generated by teachers should look like and what value there would be in collecting it (Author observation). This was partly because of the sensitivity of the issues that were related to getting data from teachers, in the context of the generally alleged economic rationalist approach and ongoing workload issues. Stehn (1997) reports that a number of related projects were underway: "Ten resource papers on issues such as programming, reporting and using student achievement information have been published" (Stehn, 1997, p 185) but in none of these was the issue of within level progress adequately considered.

The matter was also raised in ongoing evaluations of project implementation conducted by ACER in 1995 and 1996 for the SA Department (Frigo, 1997). Frigo reports that concerns identified included "the broadness of the levels, the slow movement of students through levels that don't account for progress made, consistency of levelling by teachers" (Frigo, 1997, p. 20). There was concern about the use of numbers as opposed to descriptive reporting. There were concerns about the self-esteem of older students assessed as being at lower levels than usual for a given Year level. It was alleged by one teacher that "there is no room for 'distance travelled' for the slower child, who is not ready for formal learning at 5, but may have made huge progress for his/her ability" (Frigo, 1997, p. 20). This specific criticism related to students below level 1 but the inability to record distance travelled applied at all levels. The design was intending to articulate distance travelled as one of its fundamental elements but the final product had not clarified how this might be done.

Even though some issues about progress within levels were flagged in the Frigo evaluation these did not flow into a consideration of what options there were to address them.

Data collection in South Australia

There had been a clearly stated intention to collect data from schools by 1995 (Education Department of South Australia, 1993) but this collection was delayed until 1997. In early 1997 officers of the Department for Education and Children's Services (DECS) began negotiations with a representative group of the South Australian Institute of Teachers (SAIT) about the parameters for a data collection. SAIT negotiators were clear what conditions were

to be met. Based on author memory (no documentation is available to the author) the conditions were essentially:

no teachers were to be identified at any stage,

no schools were to be identified in reports,

maximum teacher involvement with minimal time demands per teacher,

and the 'sticking point', a method to indicate progress within levels.

An existing protocol, the *Code of Conduct for Using Student Achievement Information* (1995), ensured that no students would be identified. It is important to establish at this point that these identifications related to the publication of any data that would identify an individual student or could infer identification from summary data. As early as 1993 the need to include the student identification code was made clear (Education Department of South Australia, 1993) as the link to other desired special population identifiers (gender, age, postcode, aboriginality, non-English speaking background, SES status, disability code, and Year level). Using student identification codes was an innovation of the late 1980s in the Statistics Unit, under the author's management, allowing an increase in the range of data summaries from one data collection. By 1997 it was commonplace for most statistical collections from SA government schools to be automated, using school based student records organised by student identification codes and transmitted to central office. The collection model for levels data finally adopted for 1997, and repeated in 1998, required student identification codes to enable student characteristics to be attached (age, gender, any particular sub-population identifiers, Year level). The agreed protocols ensured that these would not be published. Teachers were never considered for identification.

Negotiations for the collection broke down over the indicator of progress within levels. It was at this point the author and a colleague (Ian Probyn) were invited, as representatives of the Quality Assurance Unit of DECS, to assist in the development of collection approaches. It was unfortunate that previous attempts to discuss the approaches that might be adopted for progress within levels had not been taken up by the Curriculum Division, as it meant there were no trial approaches that had been field-tested available for consideration. The options rather hastily offered were the Victorian Model (3 Zones), a 4-zone model, a 5-zone model, and a 10-zone model. The initial meeting agreed, with surprisingly strong support from the union negotiators, to further consider a 10-zone model, once a collection process could be explained. The author and Probyn (who did the detailed development) returned about 2 weeks later with a prototype collection process.

The prototype software automatically selected the sample of students to be reported by each teacher. It had been agreed, previously, that all teachers would be involved, with a requirement of 5 students only on average to be reported per teacher. The software developed by Probyn, randomly selected one of four learning areas for each teacher. It then randomly selected 5 students for each teacher through interaction with the school's computerised student record system. For each student the strands of the selected learning area were shown on the screen. For each strand, the teacher was requested to identify the level most recently achieved and, by clicking on a continuous bar (of nine elements undifferentiated from the teacher's perspective), the progress towards achieving the criteria for the next strand indicated. The elegance of the solution to indicating progress was that it could be rescaled to any chosen divisions (2,4,5,10) of progress and did not require the teacher to consider a decimalisation of the scale, even though this was the result of a progress judgment.

The union representatives were sufficiently happy with the prototype, partly because of the automation and thus a simple response process, and partly because of the opportunity to indicate progress. They agreed that they would support a data collection of this basis. In this way the first profiles data collection was agreed to, and then carried out in 4th term, 1997. Other departmental officers who had been negotiating for some time were bemused at the ease with which the matter had been resolved. It is a matter for speculation only as to whether a more elegant process might have evolved on the issue of progress indication if more developmental work had been done over the previous three years.

The collections were conducted in Term 4 in 1997 and Term 3, 1998. The process allowed for four learning areas in each collection. After two years of data, a learning profile of the system in South Australia had been developed. A series of brochures, describing the data were provided to teachers, for each of the learning areas. These provided graphical descriptions of the trends in student progress over the year levels 1 to 8. In almost all strands in all learning areas, the general picture was the same. The median students in each Year level were on a straight line of constant gradient (about 0.4 to 0.5 of a profile level) by Year level (similar to that shown in Figure 7.2).

The distributions of student profile levels per Year level show increasing spread as Year level increases. The development of this overview of student development within strands, within learning areas, was based on teacher observation data alone. As far as the author is aware no similar data set, collected by individual student for a system-wide sample, had been

developed elsewhere[13] at that time. Rowe and Hill (1996) provide a very similar view but for a smaller sample of schools.

The general result was presented to the Australian Association for Research in Education conferences in 1998 and 1999 (Rothman, 1998, 1999). The collections were not continued following the Frigo Report of November 1998 (Frigo, 1998). In late 1998 and early 1999 the Curriculum Policy Directorate in DETE SA developed a draft writer's brief for a revised curriculum framework (Hornibrook & Wallace, 2001). Improvement of the Statements and Profiles documents was a requirement, responding to the Frigo Report (Frigo, 1998) and Withers Report of January 1999 (Withers, 1999). Initially in-house adjustment to the Statements and Profiles documents was anticipated. Following a change of Minister and Chief Executive, an outsourced revision was requested for the development of a suite of new documents and curriculum structures.

> There is no paper record available … that sources or dates the change in focus and process. However, data from interviews indicates that the change was as a result of discussions with the Chief Executive.
>
> According to data generated through the interviews, the appointment of a new Chief Executive to the department provided a different way of doing things. It was reported to the evaluation that he expected collective and connected action and expected that more minds and expertise would be brought to the task. The task was to be done in a more connected way and be done more quickly. Planning for the task was therefore mindful of the necessity to destabilise traditional working patterns and to make new connections. (Hornibrook & Wallace, 2001, p. 10)

SA moved into the next phase of curriculum reform, with the development of the South Australian Curriculum and Assessment (SACSA) Framework commencing in 1999. Eight levels were reduced to five, for the same development span, with no consideration of progress within a level (South Australian Curriculum, Standards and Accountability Framework, 2000).

Confirmation of the value of profiles – application in studies and student assessment

Meanwhile profile levels had already begun to be utilised in research. They were used in the 1996 National School English Literacy Survey, in a similar fashion to the use of the Victorian

[13] In principle the CSF 1, CSF 11 and VELS series of teacher assessment data (Victorian Auditor-General, 2009) provides a similar view. As far as can be ascertained the collection of this data in 1997 was by schools reporting means per Year level per strand rather than individual student data. Since 1998 teacher assessments have been collected electronically from schools (Department of Education Victoria, 1999).

Curriculum Standards Framework (Rowe & Hill, 1996). Masters & Forster (1997) used the level structure to 'map' the literacy skills of Year 3 and Year 5 students across the nation. The management committee for the project proclaimed the value of the approach in

> documenting the varied Levels of student achievement in those aspects of literacy which constitute the framework of the English curriculum profile: Reading, Writing, Speaking, Listening and Viewing. This is in contrast to the more limited scope of earlier national surveys of literacy achievement which were developed to gather data about the percentage of students unable to satisfy minimal levels of competence in reading comprehension. (Masters & Forster, 1997, p. iv)

The management committee also praised the value of extensive teacher judgement in the survey process.

The original profiles concept (AEC, Conference of Directors General, ACAP, CURASS) included a mechanism to assess students (and as a consequence classes and schools), directly by their teachers, on the basis of how students were making progress through the levels. Keeves (Keeves & Marjoribanks, 1999, p. 129) reports that the sequences of instruction underlying the SPFAS "have limited meaning unless there are underlying scales that would permit the assessment of student learning in the form that Masters (1982) had envisaged". Masters indicates "from an educational measurement perspective, this initiative to specify intended learning outcomes, to organise these outcomes into strands, and to describe eight levels of progress along each strand, meant that frameworks were beginning to emerge which could be used to guide test development and against which students' test performance might be reported" (Masters, 1999 in Keeves & Marjoribanks, 1999, p.289).

Embodied in Master's view of the benefits of the SPFAS was the value of the materials, not only as an explicit statement of desired outcomes from the curriculum but as a developmental ruler against which student progress could be charted. Test items and assessment tasks, while indicative of the skills achieved were of interest "only to the extent that they are useful vehicles for estimating the location of students on the variable of interest" (Masters, 1999, p. 285).

Keeves reports that the SPFAS had characteristics that "warrant the claims made of innovation, development and marked advance in a world context" (Keeves & Marjoribanks, 1999, p. 114). Features included "scales of learning and the benchmarks as levels on the scales facilitate not only teaching and instruction but also the assessment and reporting of student learning and development over time". (Keeves & Marjoribanks, 1999, p. 114). Not everyone could see the benefits of the SPFAS.

Criticisms of a level approach

While a body of pragmatic and somewhat 'cutting-edge' curriculum and assessment concepts underpinned the profiles and levels approach (Masters, 1990, 1999; Sadler, 1987; Griffin, 1990), the concept of profiles had critics from educationalists. (Reid, 1991; 1992a quoted in Jenkin, 1996; 1995)

A range of general criticisms, as well as support, appeared in the pages of Curriculum Perspectives from 1992 to 1998 (ASCA website). CURASS chair Boston (1992, 1993, 1994) explained his view on the development. Collins (1994a, 1994b) and Reid (1992b, 1995) among others, debated the merits of the national curriculum and the profiles.

South Australian based critics (Garrett & Plitz, 1999; Reid, 1991, 1992a, 1992b, 1999; Thomson, 1999; Williams, Johnson, Peters, & Cormack, 1999) were concerned about the constraints put on the curriculum by standardised approaches to design and assessment. These constraints were deemed to apply in approaches that required regular pencil and paper testing as well as in the looser and more flexible, standards approach embedded in the profiles and levels. Reid had concerns that government control, particularly national government control, would have deleterious effects, claiming "educators and school communities are shut out of decision making about the big questions" (Reid, 1999, p. 13). He believed that curriculum was an industrial issue for teachers and the then corporate pressures were to increase control, to define the 'good' teacher as a "skilled technician who can most effectively deliver the expectations set by the curriculum" (Reid, 1999, p. 192) with little professional autonomy and little school autonomy. This concern seems partly at variance with the strongly declared intention of the profile concept to empower teachers, rather than tests, as the adjudicators of student progress (Boomer, cited in Jenkin, 1996).

While other commentators (Lokan, 1997; McGaw, 1994) saw more utility to the design, at least as an assessment support, it is not surprising that there was uncertainty about the classroom use of levels in the minds of teachers.

Critics of 'outcomes based' curriculum approaches added to the complex challenges and possible confusion facing teachers. Donnelly (2007) insists that the Outcome Based Education (OBE) approach followed in the development of the SPFAS was strongly influenced by Spady (1993). Spady was a US advocate of a range of outcomes approaches (Donnelly, 2007, p. 2). Spady, however, is not referenced directly in background documentation as far as the author can ascertain, although the Australia Curriculum Studies Association sponsored a tour by Spady in late 1992 and published his material (Spady, 1993). Certainly the SPFAS was dependent upon clear descriptions of outcomes. Spady's considerations of outcomes may have influenced some teachers and some planners but they

were not the main influences upon the design. As indicated earlier, locally developed concepts (Griffin, 1990; Masters, 1990; Sadler, 1987) were much more powerful influences.

While the profiles were evolving and replacing the attainment levels, South Australia, under the direction of the Brown Liberal government elected in 1993, was also implementing a statewide testing program.

A Parallel Universe - the Testing Approach

A separate initiative was developed in South Australia in the period from 1994 to 1996 and has been repeated annually since then. From 2008 it became part of the *National Assessment Program-Literacy and Numeracy* (*NAPLAN*). This was the introduction of a testing program known as the Basic Skills Testing Program (BSTP) in 1995, after a trial in 41 schools in 1994.

Over a decade earlier Keeves (1982), chair of the Committee of Enquiry into Education in South Australia, had recommended the introduction of revised approaches to student assessment. A concern about student assessment raised by Keeves was the "apparent time consuming nature of such activities in the classroom" (Keeves, 1982, p. 184). The committee argued the benefits of observation schedules among other possibilities to increase the range of skills being assessed. In addition the committee recommended that schools be encouraged "to conduct each year a testing program in the areas of essential skills, numeracy, oracy, reference, problem solving and investigations at the Year 5 and Year 9 levels" (Keeves, 1982, p. 187) but with the decision to implement the assessment to be a local decision for each school. No action was taken to implement the recommendation.

Thus at the point where South Australia was implementing the SPFAS, a parallel development to test students at primary level was introduced, adding to the mix of industrial tension. To introduce a statewide test, South Australia contracted the NSW Department of Education to develop and mark the test in the initial years. The test was conducted at Year 3 and Year 5, and extended to Year 7 in 2001. The declared major purpose of the BSTP was to identify students having difficulties in areas of numeracy and literacy. Each participant was given an individual report indicating items correct and incorrect, graphed in difficulty order, as part of an individual diagnostic analysis. Based on test performance the student was allocated to one of 6 band levels (Hungi, 2003). The initial data collections from 1994 through to 2000 have been extensively analysed by Hungi (2003). The 1997, 1998, 2001 and 2002 waves of the data are part of the analysis reported in this thesis.

The introduction of the test program was controversial. Major concerns were expressed by teachers, while many parents and politicians supported the test program. Hungi (2003) summarises the main arguments. Critics saw the program as unnecessary, not superior to

teacher assessments, and likely to cause teachers to alter their classroom instruction to match the tests and neglect other parts of the curriculum. Supporters saw the program as providing useful feedback, particularly in identifying and assisting weaker students. The program would also assure parents of the quality of the programs in public schools. In 1998 over 95% of the target populations participated in the test program, comparable to other testing years, and confirming for Hungi that parents in the main supported the program (Hungi, 2003, p. 3).

Hungi researched the BSTP data to test the fit of the data to the Rasch model, to explore the item difficulties in Years 3 and 5 and to adjudge whether a common scale could be used for items in both literacy and numeracy for the period 1995 to 2000. He was also interested in the changes in the performance of cohorts of students over time and whether he could quantify growth from Years 3 to Year 5 (Hungi, 2003, p. 6). Hungi's analysis is very detailed and comprehensive. He reports that

> overwhelmingly, the items had adequate fit to the Rasch model and the item means between Grades 3 and 5 compare well year after year. Clearly, the test developers did excellent work in the development of the items and in the allocation of the items to either the Grade 3 or Grade 5 tests. (Hungi, 2003, p. 107)

He also established that the growth in achievement between Years 3 and Year 5 for both numeracy and literacy was consistently about 0.50 logits per year for each of the 6 years in the analysis. He remarks (p 107) that this growth has consistently increased slightly each year, "especially for numeracy". Hungi's detailed analysis of the Basic Skills Test data, re-used in this thesis, provides evidence that the tests are of high quality, fit the Rasch model well, and provide quality reference measures for comparison with teacher's judgements of students involved in both assessment processes. The test data are summarised in Chapter 6.

Summary

The chapter has outlined the general history of the development of the profiles approach adopted in South Australia in the mid 1990s. The model for assessment of student progress was taken directly from the SPFAS, although a parallel attainment levels approach had been developed immediately prior to SPFAS. While the model was implemented in the face of considerable opposition from teachers, this opposition was more about perceived workload than fundamental objections to the assessment model. At the same time as profiles were implemented in SA, primary teachers were objecting to the introduction of tests at Years 3 and 5. A brief synopsis of the beginning of statewide testing was also provided.

Initially, a data collection of teacher assessments of students using the profiles was expected for 1995. A final process was not resolved until mid 1997. Little interest was shown in the refinement of assessments to establish progress within a level. The issue of finer resolution

within a level came to a head in the planning for a collection of student data when teachers refused to provide data unless it was at a finer resolution than a level. A collection process was designed and applied without trial that enabled teachers to indicate the progress a student was making towards achieving the next level beyond the one the teacher believed the student had already achieved. Those data provided by teachers in 1997 and 1998 are analysed in this thesis.

Sadler's concept of standards-based assessment using a process of teacher judgement was a key element of the SPFAS. The next chapter reviews the broad issue of teacher judgement and cites cases from the research literature that illustrate how it has been applied and how it compares to independently obtained measures of student learning.