



Accountability: Inclusive assessment, monitoring and reporting



AUTHORS

Allan Luke, Annette Woods, Ray Land,
Mark Bahr & Margaret McFarland

School of Education
The University Of Queensland



Research Report prepared for the Indigenous Education Consultative Body, Brisbane

ACCOUNTABILITY: INCLUSIVE ASSESSMENT, MONITORING AND REPORTING

**Allan Luke, Annette Woods, Ray Land,
Mark Bahr & Margaret McFarland**

**School of Education
University of Queensland**

**Research Report prepared for the
Queensland Indigenous Education Consultative Body**

ACKNOWLEDGEMENTS

The authors wish to thank the members of the IECB for their commitment, encouragement and direction. The project would not have been possible without the expert guidance of Patricia Stroud. Staff from the Queensland School Curriculum Council and Education Queensland provided documentation, expert opinion and data. Michael Byrne of Education Queensland generously provided time and expert technical advice. Shane Williams of Education Queensland has lent moral support and direction.

The report relied heavily on input from an expert panel of senior Aboriginal and Torres Strait Islander educators: Lloyd Appo (Bundaberg West State School), Will Davis (Beenleigh State High School), Steve Foster (Badu Island State School), Maureen Liddy (Hopevale State School), Martin Nakata (University of South Australia), Chris Sara (Cherbourg State School) and Grace Sara (Cherbourg State School).

At University of Queensland, Jodie Gaven and Rod Neilsen provided research assistance during the project; Professor Peter Renshaw, Griffith University offered conceptual input in the initial phases of the project.

Finally, we wish to thank the administration and staff at the state, catholic and independent schools who offered advice, documentation and access to their, and their students' experiences and practices. We hope that this research makes a difference for them.

EXECUTIVE SUMMARY

SUMMARY

This aim of this study was to investigate how current practices for assessment, evaluation and reporting of student outcomes in Queensland schools address distinctive educational needs, challenges and capabilities of Aboriginal and Torres Strait Islander children. The study describes the contexts and current status of a broad range of assessment practices found in school – informal and formal, standardised and non-standardised. These include: system-wide standardised testing, school-based diagnostic procedures, and classroom based evaluation and reporting to parents. The major focus of this study is on a meta-analysis of the 2000 Queensland Year 5 Aspects of Literacy and Numeracy Tests results, items and data sets. Beyond the publicly available test score results, this was the most comprehensive and current data set made available to the research team. The analysis of this data was used to identify several key areas for further research, development and policy. The report concludes with recommendations for system and school-level interventions in the area of testing and assessment.

RESEARCH METHODS AND PROCEDURES

The study involved five major steps, conducted between October 2001 and April 2002:

1. A review of literature on ‘culture-fair’ assessment, test bias and discrimination and literature on assessment and testing of Indigenous children. This review was used to develop a template for looking at the problematics of bias and fairness in large-scale standardised testing.
2. Discussions with representatives of the Queensland School Curriculum Council and Education Queensland, and analysis of existing documentation on tests. This was used to document the policies and development processes that led to the current Year 2 Diagnostic Net (Net) and Queensland Year 3, 5 and 7 Tests in Aspects of Literacy and Numeracy (tests).
3. Field visits to a sample of four schools to provide a broad range of assessment, evaluation and reporting practices used with Indigenous children. This was used to sample local school and classroom uses of Net and test data, other school-level approaches to assessment, and to examine how test and other performance data is reported to Indigenous parents and families.
4. Convening of an expert panel of Aboriginal and Torres Strait Islander teachers, principals and researchers to undertake a content analysis of selected standardised tests to identify specific items for linguistic and statistical reanalysis. This involved an item-by-item discussion and expert rating of the 2000 Year 5 Aspects of Literacy and Numeracy Tests for possible linguistic

and cultural bias. Linguistic and content analysis, and statistical re-analysis of Indigenous student performance on specific items identified by the expert panel.

5. Statistical reanalysis of overall 2000 Year 5 test results by population, background and location.

FINDINGS

- F1.** That there is some evidence that some specific test items discriminate against Aboriginal and Torres Strait Islander students on the basis of culture-specific background knowledge and linguistic background.
- F2.** That there is some evidence that the test administration and reporting formats discriminate against Aboriginal and Torres Strait Islander students on the basis of culture-specific background knowledge and linguistic background.
- F3.** That efforts to extend tests to cover more complex syllabus outcomes, constructivist problem solving in numeracy, and other areas had generated a technical complexity, as well as complex issues of cognitive load and construct validity for the structure of the tests.
- F4.** That there have been repeated documented calls for a critical re-evaluation of the cultural-specificity of the First Steps Developmental Continua as used on the Year 2 Diagnostic Net, and its affiliated support materials.
- F5.** Although there was some evidence in the small sample of good school-level practice, the research found that the actual school and classroom based practices for assessing students and reporting on their results tend to be, at best, highly variable and school-specific, and at worst, *ad hoc*.
- F6.** There was highly uneven evidence that system authorities had recognised that there might be problems with assessment bias and that system level attempts to adjust, accommodate or make more appropriate existing instrumentation and materials were sporadic, under-funded and had not had a major impact on practices in the field.

RECOMMENDATIONS

- R1.** That the QSA, EQ, QIECB and relevant systemic authorities should establish a working group on assessment of Aboriginal and Torres Strait Islander students, with requisite, appropriately funded technical expertise in measurement *and* cultural issues. This group's charge needs to go beyond consultation and 'sign-off' on assessment instrumentation to involve the development and oversight of a proactive research and development agenda to put Queensland in the vanguard of culture-fair assessment.

- R2.** That specific achievement data should only be used for the purposes it was designed to be used for. That information about the purposes, limits, uses and abuses of specific Net and test data be circulated widely across all systems, and that critical caveats on the misuse and misinterpretation of achievement data for purposes it was not designed for be put in place. For example, non-generalisable data should not be generalised. Individualised data should not be disaggregated. Formative and diagnostic data should not be used for summative purposes. There is extensive evidence across the system of sanctioned abuse and misrepresentation of Net and test data.
- R3.** That teachers and schools working with Indigenous students should receive both system-wide advice and training in approaches to culture-fair assessment, on techniques for interpreting test data and its limits, and on techniques and approaches for reporting on student performance to parents, Elders and communities in meaningful and educationally useful ways.
- R4.** That QSA, EQ and other systems develop and trial culturally relevant materials for Aboriginal and Torres Strait Islander students, for use in face-to-face diagnostic assessment (e.g., the Net and Reading Recovery Observation Survey)
- R5.** That QSA, EQ and other systems undertake a study of whether the Developmental Continua and the outcomes statements that relate to levels 1, 2 and 3 of the Years 1-10 English Syllabus presently being trailed, are able to accurately accommodate multiple and diverse sequences of the development of literate practices.
- R6.** That there is a need, as an experimental research and development task, to prototype and trial several alternative test items on Year 3, 5 and 7 tests that accommodate and use Indigenous knowledge.
- R7.** That the possibility of developing and deploying alternative and adjusted test administration environments and procedures for children in remote or EFL communities should be examined.
- R8.** That there is a need to re-examine the language demands and construct validity issues in the Years 3, 5 and 7 numeracy tests.
- R9.** That there should be an effort to develop simpler, less linguistically contingent and technically complex formats and item structures of assessing basic numeracy and literacy in future tests. This may require a reconsideration of efforts to extend syllabus coverage through the tests.
- R10.** That the QSA, EQ and other bodies should examine alternative categories for internal reporting and analysis of state testing results, disaggregating Aboriginal and Torres Strait Islander achievement by location, by main language experience (EFL/ESL/ESD) in order to better identify ostensive gains and losses among some sub-groups of students.

TABLE OF CONTENTS

| | |
|---|----|
| Coverage of Key Outputs for the Tender in Report | 1 |
| 1: Introduction | 3 |
| 1A. The Problem and the Questions | 3 |
| 1B. The Purposes and Functions of Evaluation | 5 |
| 2: Bias in Assessment and Monitoring | 10 |
| 2A. On Culture Fair Testing | 10 |
| 2B. Complex Issues of Content Validity | 12 |
| 3: Indigenous Education and Assessment Policy and Practice | 15 |
| 3A. Indigenous Education and the Deficit Mentality | 15 |
| 3B. The National Move to Outcomes Based Education | 17 |
| 3C. The Queensland Year 2 Diagnostic Net and Indigenous Learners | 18 |
| 3D. The Queensland Year 3, 5 and 7 Testing Program and Indigenous Learners | 26 |
| 4: 2000 Queensland Year 5 Tests in Aspects of Literacy and Numeracy - A Case Study | 29 |
| 4A. Indigenous Educators' Panel | 29 |
| 4B. A Linguistic Perspective | 32 |
| 4C. A Quantitative Perspective. | 37 |
| 4D. Summary | 49 |
| 5: Policy Analysis Issues for Applications of State-Wide Accountability Measures | 51 |
| 5A. Positive By-Products of the Year 2 Diagnostic Net Process | 51 |
| 5B. Limitations of The State-Wide Tests | 53 |
| 5C. Institutionalisation of the Net and Tests as Systemic Accountability Measures | 54 |
| 6: Conclusion: A Clear Way Forward | 57 |
| 6A. Findings | 58 |
| 6B. Recommendations | 59 |
| References | 61 |
| Appendix 1: Case Studies of Effective Practices | 67 |

LIST OF TABLES

- Table 1:** Discriminant function analysis summary for the four literacy sub-domains of the 2000 Year 5 Aspects of Literacy Test
- Table 2:** Analysis of reading and viewing items from the 2000 Year 5 Aspects of Literacy Test sorted by discriminating power
- Table 3:** 2000 Year 5 Aspects of Literacy Test Item 15 response patterns broken down by indigeneity
- Table 4:** 2000 Year 5 Aspects of Literacy Test Item 17 response patterns broken down by indigeneity
- Table 5:** 2000 Year 5 Aspects of Literacy Test Items 19 - 21 response patterns broken down by indigeneity
- Table 6:** 2000 Year 5 Aspects of Literacy Test Item 29 response patterns broken down by indigeneity
- Table 7:** 2000 Year 5 Aspects of Literacy Test Item 35 response patterns broken down by indigeneity
- Table 8:** 2000 Year 5 Aspects of Literacy Test Item 35 Indigenous students' response patterns broken down by NESB
- Table 9:** 2000 Year 5 Aspects of Literacy Test Item 35 non-Indigenous students' response patterns broken down by NESB
- Table 10:** 2000 Year 5 Aspects of Literacy Test Item 35 raw performance by district
- Table 11:** 2000 Year 5 Aspects of Numeracy Test - Discriminant function analysis summary
- Table 12:** 2000 Year 5 Aspects of Numeracy Test Item 14 response patterns broken down by indigeneity

LIST OF FIGURES

- Figure 1:** 2000 Year 5 Tests in Aspects of Literacy and Numeracy patterns of variation of response by indigeneity and gender

1: INTRODUCTION

1A. THE PROBLEM AND THE QUESTIONS

Despite a continued focus and substantial effort by Australian governments at all levels, educational inequality for Aboriginal and Torres Strait Islander peoples persists at intolerable levels. It is possible to measure and assess the scope and scale of educational inequality on qualitative *and* quantitative indicators: both by biographical and oral, narrative and literary accounts of Aboriginal and Torres Strait Islander Elders and children, and by a range of systems-based performance indicators. The latter are the focus of this research.

Educational equality and inequality, effectiveness and ineffectiveness, success and failure can be assessed by a range of institutional and social outcomes indicators. These include, but are not limited to: attendance rates, retention and attrition rates, rates of behaviour management incidents, parental and student satisfaction survey data, teacher absenteeism and tertiary success. But under current orientations to educational policy and governance, systems routinely collect and analyse a range of *de facto* measures of academic achievement, performance and outcomes. These include: teacher and school-assigned grades and marks, face-to-face diagnostic assessments, and secondary matriculation and or graduation results as assessed and allocated by various state systems.

Compared with the USA, UK and Canada, Australia does not have a long history of universal standardised achievement testing (Luke & van Kraayenoord, 1998). However, since the early 1990s, State, Territory and Commonwealth governments have begun to move towards the assessment of educational equality and inequality, system efficacy and accountability by reference to differential patterns of achievement by various 'target groups' on norm-referenced, standardised achievement tests. As will be described here, this has involved the development of state-wide testing processes mandated for government schools and optional for non-government schools, and a complex process of setting national 'benchmarks' or criterion referenced performance standards between state tests. In this context, the role of assessment and testing is ethically, scientifically and politically significant. This is particularly the case with reference to the education of Aboriginal and Torres Strait Islander students, which faces a period where Indigenous communities, governments and others are demanding demonstrated evidence of improvement in provision, intervention and outcomes.

The picture of Aboriginal and Torres Strait Islander performance on these instruments is not a positive one. In the 1996 National School English Literacy Survey, approximately 70% of students in Year 3 were reported as having met the identified performance standards in reading and writing. This overall result compares with the results that only 20% of Indigenous students at this level met the reading standard, and less than 30% of this group met the writing standard (Commonwealth Government of Australia, 2000, p.9).

In their 1998 ACER study of longitudinal change in middle years reading comprehension achievement, Marks and Ainley (1997) reported an overall improvement over a ten year period of Aboriginal and Torres Strait Islander students' performance. However, these students still lagged behind both NESB and non-Indigenous students' overall results. There is some evidence in the 2000 National Report on Schooling (DEST, 2000) that the overall reading achievement performance in Years 3 and 5 of Aboriginal and Torres Strait Islander students in Queensland and nationally has improved marginally between 1999 and 2000. Yet despite evidence of some degrees of overall progress, Indigenous children in all states still achieve at significantly lower levels than non-Indigenous Australian children, including NESB and immigrant children.

In Queensland, Aboriginal and Torres Strait Islander (ATSI) students make up a significant but small proportion of the overall student population, however they figure disproportionately in the numbers of students performing at lower levels on state-wide assessments, and thereby those requiring additional support. Further, there are considerable variations in performance levels within the Indigenous student group - with rural students, especially those living in remote/community settings, performing significantly below their urban cousins. There is also evidence that the gap between Indigenous and non-Indigenous students widens as cohorts progress through their primary schooling into the secondary years, particularly for male students. Several examples may make the relative sizes of these patterns of disadvantage clearer.

While ATSI students made up just 6.8% of the Year 2 population in 1998, the gap between the Indigenous and non-Indigenous students being identified as needing support was 28.5% for reading, 28.7% for writing and 32.2% for number. While the comparable gap percentages for 2000 demonstrate some overall improvement for Indigenous students (17.2%, 19.8% and 10.6% respectively) this performance gap remains the most student group significant difference apparent in the data. Further, a rural Indigenous student remains one and one-half times more likely than his or her urban cousin to be identified in Phase A as requiring of additional support for reading, writing and number, suggesting that performance gains for Indigenous students over recent years have largely been confined to students attending urban schools. (For current data, see <http://education.qld.gov.au/schools/statistics/html/studstst.html>).

Similarly ATSI students made up 6% of the Year 5 population in 1998, but the gap between Indigenous and non-Indigenous students identified in the lowest 15% of achievers in both literacy and numeracy was 30% or greater (Performance Measurement and Review 1998, cited in Education Queensland 2000a, p.22). The statistical data available at the EQ website identified above strongly suggests that despite the identification of such literacy and numeracy difficulties in Year 5 or earlier, the overall gap between Indigenous and non-Indigenous students continues, if not widens, by the Year 7 Test and subsequently to secondary schooling. While the overall percentages of ATSI students had dropped to 4.1% of the Year 10 population in 2000 Indigenous students, particularly males in rural secondary schools, are still far more likely to be judged as 'failing' by the end of compulsory schooling – that is, to be disproportionately accorded Low or Very Low Achievement levels in subjects such as English and Mathematics.

The continuation of such patterns of disadvantage has clear policy implications for governments, as well as personal life consequences for young Indigenous people. The Commonwealth Government reports that while access and participation rates are improving, equitable education outcomes are still a long way from being achieved. The Commonwealth argues that to achieve such will require a co-ordinated national approach that addresses all issues related to educational success (Commonwealth Government of Australia, 2000, pp.8-9).

For Indigenous communities, education systems and researchers alike to use poor test and Net scores as evidence of unfinished educational business and, conversely, to cite improved test and Net scores as evidence of improvement, success and equity begs a set of prior analytic, empirical and technical questions.

The questions for Aboriginal and Torres Strait Islander parents and communities, Elders and educators are:

- What does this data say to us?
- What do these instruments assess?
- Do they assess and report in ways that are relevant, accurate and fair?
- And how might we develop, use and interpret assessment instruments and techniques that will work in the educational interests of Indigenous children and communities?

These are the key questions that underlie this report.

1B. THE PURPOSES AND FUNCTIONS OF EVALUATION

Assessment and testing serve major, essential functions within education systems: as tools for the organisation of teaching and learning (Masters & Mislevy, 1993), as means for accountability for teachers, schools and jurisdictions, and as gatekeepers for entry to education programs of all types. No modes of assessment are completely objective, truthful and scientifically infallible. All can be subject to and mediated by cultural ideology and bias. For example, a teacher-constructed test to check that a concept has been taught correctly can misrecognise or misconstrue student cultural or linguistic competence and knowledge. An external marker or examiner can 'embody' particular values and dispositions in the assessment of a student essay or project. To counter assessment bias, to make assessment at all levels as fair as possible, requires a knowledge of how tests work, how they are constructed, and what they can be reasonably used for.

All standardised testing begins from three baseline assumptions. These are that:

- people can be ranked;
- whatever is being measured is intrinsic to the individual (and not the culture or community of which the individual may hold membership);
- whatever is being measured actually exists as a distinct and measurable entity.

While this may seem straightforward – the making of fair, accurate and valid tests is technically challenging. Standardised testing continues to be informed and directed by mainstream, dominant values, beliefs and cultural practices (Strickland 2000). Its

function, broadly in legitimating and streaming access to interventions, pathways and resources, is necessarily tied to the beliefs and assumptions of dominant cultural and social class groups – those whose values, knowledges and practices are actually coded and reproduced in schools and other educational institutions.

It is within this context that the question of whether standardised achievement measures used in English speaking and mainstream educational systems can be made less discriminatory against students from Aboriginal and Torres Strait Islander cultural and linguistic backgrounds needs to be asked. Many researchers argue that such efforts are futile, that by definition testing in mainstream institutions functions to discriminate. That all measures of skill and cognition are intrinsically culturally-based in definition, procedure and structure. By this account, the idea of somehow making them free of discriminatory bias is theoretically and practically unachievable (e.g., Blum, 1978), particularly in educational institutions with unenviable historical track records on racial and ethnic discrimination (Berlak, 2001). Tests and testing cannot be culture free.

By contrast, there are recent studies that suggest that, while cognition and skill are culturally-based, assessment tasks can be tailored in their content, format and approach to better assess what members of cultural groups actually can do, cognitively, behaviourally and academically (e.g. Scribner & Cole, 1981) In Scribner and Cole's milestone study of literacy and cognition among the West African Vai tribes people in the late 1970s, they demonstrated that classic Piagetian tasks, when reframed in culturally comprehensible, everyday contexts (e.g., using locally available examples, implements and tools) generated better performance among this group, than did the use of traditional Western assessment procedures. In some of the US literature cited below, there is evidence that when tests are adjusted in cultural format and content, the overall performance of non-mainstream ethnic groups can improve. So the issue then is to work toward making tests and assessment procedures as culturally fair as possible.

The concept of culture-fair assessment and evaluation of students has a long history. It is certain that classroom, school based and systems-based evaluation of student performance, outcomes and work must be by definition a cultural construct. That is, it must reflect the normative values, purposes and procedures of specific fields of knowledge which are typically Western and Eurocentric. Further, it reflects by necessity specific contexts where behaviourally and cognitively assessable skills are used. These contexts are also usually non-Indigenous in nature. Finally educational evaluation must reflect schooling and education more generally, and in Australia this remains very much based in mainstream Anglo-European cultural practices and traditions. It also reflects the histories, precedents and traditions of assessment. These include the examination, which historically can be traced back several hundred years to Chinese civil service examinations, Catholic and Hebraic textual initiations and examinations, through to the American standardised achievement test, which can itself be traced back to the work of Binet, Weschler and others and the early 20th century 'sciences' of education. All of these traditions of assessment are central to the way schools and other educational institutions work, and all potentially provide differential access to competence, success and knowledge.

What this means is that all assessment and evaluation by definition reflects particular values, cultures, traditions and ideologies. This includes teacher judgement in examination marking and classroom evaluation, which is as potentially discriminatory on class, race and ethnicity, language and gender grounds as standardised assessments. In their studies of teacher-based examination marking, Freebody (1990) and earlier Ozolins (1981) both found that teacher-judgement may subtly search out and assess linguistic markers of culture and ethnicity, social class, language and gender. If assessment and evaluation entails assessing student performance vis a vis 'norms', benchmarks, standards or even well defined domains of performance – they cannot be culturally neutral.

There are, then, three key questions facing Indigenous communities, educators, parents and families and Elders:

- **Content validity:** Are the cultural practices, norms and values embedded in assessment - that is the actual knowledges, skills and competencies - that the tests purport to test, the relevant and valued outcomes that these stakeholders seek from education?
- **Construct validity:** Do the tests and other diagnostic scales actually test what they purport to test? and
- **Reliability and bias:** Do the tests operate fairly, without discriminating against the performance of any particular group or subgroup being assessed?

What is certain is that there must be forms of assessment, evaluation and measurement in Queensland schools. It is clear to us that there are powerful educational imperatives for assessment and evaluation, which can serve distinctive purposes. These include:

- **Diagnostic and formative assessment of individuals** which can guide teacher and school selection and framing of intervention, and assist in reporting student progress to parents;
- **Formative evaluation of programs** which can guide system, school and teacher selection and modification of curriculum and instruction;
- **Summative evaluation of students** used to identify select students for pathways to, through and across educational systems; and
- **Summative evaluation of school, program and system efficacy** which can guide systemic allocation of resources, systems and community accountability.

Given the current demands among Aboriginal and Torres Strait Islander Elders, educators and communities for improved educational outcomes as a key part of an overall economic and social strategy, there is a further purpose for educational assessment: to inform communities, consultative and advisory bodies like the QIECB, and their affiliated constituencies, of the relative success and failure of schools in improving students' outcomes and life chances.

A persistent problem in the educational literature is matching assessment techniques and instruments with their proper and suited educational functions. That is, given the wide range of specialised forms of assessment and evaluation, it is imperative that teachers and systems use test and classroom evaluation instrumentation for the purposes for which those instruments were designed. *Unfortunately, at present, the Year 2 Diagnostic Net, and the Year 3, 5 and 7 Tests in Aspects of Literacy and Numeracy are being used for all of the above mentioned functions, often without systemic guidance, caveats or clarifications on their use.*

We began this research project from the baseline argument that testing or assessment cannot be ‘culture-free’, or uncontaminated with the values, beliefs, and culture of those planning, designing and using the assessments. Whether we are dealing with Queensland’s approach to state-wide literacy and numeracy testing, the Year 2 Diagnostic Net procedures, the portfolio-based, teacher judgement system used by the Board of Senior Secondary School Studies, or TAFE/vocational education competence assessment – these assessment and monitoring systems by definition are codings of particular workplace and school cultures, and specific cultural values and ideologies. Systems and schools must assess, and they must assess in relation to particular normative curricular goals, skill outcomes and credentialing standards.

At this historical juncture, as is reported in this study and a concurrent study on ESL and language and literacy teaching in Indigenous Communities (Luke, Land, Christie & Kolatsis, 2002), we found a strong consensus amongst Indigenous educators and community members, among educators who work with Indigenous students, and among researchers and bureaucrats working with assessment and curriculum. All were committed to raising the ‘standards’ and the overall achievement, performance and outcomes of Indigenous students. All were committed to improving the Standard Australian English literacy of Indigenous students. And all parties are tired of ‘excuses’, of rationalisations over poor performance. All parties understand that raising these standards will require assessment – and that the assessment must rigorously pursue standards and outcomes of mainstream skill, knowledge and competence.

However, all also wanted to see that the instruments, techniques and procedures used for these purposes do not systematically disadvantage children from Indigenous cultures and communities or preclude rather than enhance their subsequent engagement with mainstream cultures and knowledges. Stakeholders across Queensland want as fair and as equitable a playing field as is possible and they want those who develop and work with assessment instrumentation to remain ever vigilant against the potentially powerful discriminatory effects and consequences of testing and assessment.

As the subsequent overviews of the field will show – the history of assessment of Indigenous peoples and cultural ‘minorities’ is not a pretty one. For Indigenous Australians, it had its basis in the first Cambridge scientific expeditions which tested, measured and assessed every physiological, psychological and behavioural aspect of Torres Strait Islanders at the beginning of the last century. Suffice to say that the normative instruments and measures used were based on European studies of colonised peoples across the globe (Nakata, 1997) and were thus unable to assess competence in any measure without reference to Western assumptions and standards.

Part of the problem of these early 20th century ‘scientific’ measures was their treatment of all ‘indigenes’ as of a kind of homogenous and singular group with particular measurable racial and ethnic characteristics. This tradition continues within the current Queensland accountability measures in literacy and numeracy. The reporting of data by racial, ethnic category allows Aboriginal and Torres Strait Islander peoples to again be categorised as a homogenous group. It may thus hide differential patterns of success and failure within and across Indigenous communities, patterns of which would be vital for our understanding of both disadvantage and effective interventions.

The question for us and for this report is whether given the diversity of students in Australian and Queensland systems – with over 20,000 Aboriginal and Torres Strait Islander students enrolled in government and non-government schools, and estimates of NESB and ESL speakers ranging from 16-20% - whether the assessment approaches and instruments, and their particular cultural values, traditions and practices, take into consideration this diversity and manage to accommodate and cater for it without compromising the tests’ scientific integrity, accuracy and educational purposes. It is in the interests of all parties that these assessment instruments give us accurate pictures of what they purport to assess, and that they enable Indigenous students to show their optimal educational achievements, as they would and should enable children from mainstream backgrounds to do. *It is our belief, on the basis of the research represented here, that the answers can be had through simple, basic assessment tools with clearly defined purposes, domains, limits and uses.*

The research reported here, then, is a deliberate attempt to act on behalf of the educational interests of Indigenous communities and children. Our starting point, as we have stated in this introduction, is that assessment and evaluation is a necessary component of education systems. It is also a key strategic necessity if we are to improve Aboriginal and Torres Strait Islander education and outcomes. If this is the case, then the assessment must be as rigorous and accurate, as fair and as equitable as possible.

2: BIAS IN ASSESSMENT AND MONITORING

2A. ON CULTURE FAIR TESTING

Background

The first formal psychometric tests were face-to-face task inventories used to identify children with intellectual impairment in France, England, Germany and the United States over 100 years ago. By the end of World War I, American educational researchers had developed large-scale 'pencil and paper' tests – moving from face-to-face inventories to formats that could be administered in economies of scale to large population groups (Gould, 1981). These tests became central to streaming, placement and certification of students in American schools, the widespread adoption of pre-test/treatment/post-test approaches to curriculum and instruction, and the use of test scores to quantify educational "outputs" (Callahan, 1962).

The original purposes of standardised tests were to identify those whose ability or performance were so different from the norm on some dimension as to justify specific educational intervention and support. It is not surprising, then, that the very construct of norm-referencing leads directly to concepts of deficit, where difference in performance is interpreted as 'lack' or 'deficiency' of skill, competence or knowledge of the mainstream.

Questions of whether standardised testing was 'culture-fair' were first raised in the 1960s, in the aftermath of American Civil Rights legislation. Then African-American groups raised a series of legal test cases that argued that educational placement, employment and other judgements made on the basis of test scores were discriminatory (Blum, 1978). Until that time, American education systems had been characterised by the widespread use of standardised, norm-referenced achievement testing in two major areas: intelligence testing and basic skills testing. The results of these tests had also given rise to claims that particular groups' (specifically African-Americans) consistently poor performance on such measures were the results of, variously, innate lacks in intellectual capacity skill and capacity 'deficit' generated by community cultures (Rose, Kamin and Lewontin, 1984).

Although much of the early 20th century development of testing focused on the identification of children with disabilities, as per Binet's original intents for the identification of giftedness - some researchers (see for example Livingston 1995, Strickland 2000) have traced the historical link between standardised testing and discourses of ethnocentrism and racial superiority. Since the 1970s, there have been persistent debates and attempts to develop models of culture-fair assessment, based on attempts to identify how cultural difference influences and impacts upon test performance. In the US, Black-White differences of approximately 1 standard deviation and Hispanic-White differences of approximately 0.6 -0.8 standard

deviations have been widely and consistently reported for measures of cognitive ability (Hartigan & Wigdor, 1989; Sackett & Wilk, 1994).

The debates over innate intellectual capacity by race and ethnic background, over conflicting definitions of intelligence, and over the educational effects of intelligence testing continue to this day. They were recently raised again in the 1990s US debate over the “bell curve” (see as examples: Fischer, Hout, Jankowski, Lucas, Swindler & Voss, 1996; Herrnstein & Murray 1994, Jacoby & Glauberman, 1995). Although intelligence tests are still used by educational psychologists in Queensland to assess learning problems, the tests under question in this report are not intended to measure intelligence per se. Rather they focus on skill and knowledge-related constructs (e.g., specific literacy or numeracy skills or subskills.).

Recent work in the US by Berlak (2001) argues that, even when economic factors are controlled, African-Americans, Native Americans, Latinos and immigrants for whom English is not a first language lag behind English speaking white students in achievement in school knowledge. This is the case even in areas that are not overtly language-related in content and domain definition. Berlak argues that students from a non-dominant culture experience testing as a form of cultural intimidation. Students from a culture perceived as having limited ability face a series of stereotypes not encountered by other students. What he terms as “stereotype vulnerability” (p.3) may be created by a pattern of resistance to the values of the dominant culture, characterized by ambivalence about the importance of academic success. Berlak’s key point is that students from particular ethnic and racial groups may actually develop attitudes and practices of resistance to the surveillance, judgement and categorisation practices that are affiliated with large-scale testing.¹

What is Culture Fair Testing?

We now move to discussing what culture-fair testing may be taken to be. To begin with, it is not an attempt to renorm or statistically manipulate results in favour of the culturally different group. To define culture-fair assessment, we need to reappraise how different cultural practices may ‘interfere’ with optimal performance of particular groups in standardised test environments. The literature suggests a range of variables that may influence test performance. These include:

The cultural-specificity of how the task or activity in question is framed: As Glick (1974, in Ross, 1984) pointed out, some cultural groups approach tasks and behaviours with a different sense of boundary, speed and time requisite for solving the problem. In her early Australian work on socio-cultural aspects of cognition, Goodnow (1976) argued that Western tasks often place a premium on solutions that do not involve manipulation of test materials (see also M. Cole, 1999). Such tasks, it should be added, also presuppose particular attitudes about the authoritative nature of

¹ We note that this could form the basis of an interesting research study which has not up until this time been undertaken with Aboriginal and Torres Strait Islander children. This being a survey and documentation of the attitudes, beliefs and responses of these students to ‘testing’, its practices, and all that accountability stands for.

text-based and teacher knowledge in educational settings and how to interact with the same.

The cultural-specificity of the normative models of child and adolescent development reflected in the domain specification and constructs of the test

(Ross, 1984): Specifically, it may be that, the expected developmental sequence of competence, skill or behaviour anticipated in the instrument may be at variance with culture-specific development of skill and competence. This is particularly the case in relation to literacy development, where there is extensive literature that demonstrates multiple culture and language-specific pathways to literacy development (Gregory, 1997).

The linguistic codes and conventions of the test and task: It is axiomatic in language and discourse studies that language use, convention and knowledge shapes the categorical way that we experience and see the world and that different languages and their affiliated cultures have distinctive lexical-grammatical resources for solving problems and shaping the world (Halliday, 1978).

The cultural-specificity of content knowledge: Specifically, if learning is founded on schemata, and structured background knowledge tends to be culture-based (van Dijk & Kintsch, 1983), then any assessment of knowledge is contingent upon access to and prior experience with particular culture-specific background knowledge. Where aspects of social common sense or 'social intelligence' are brought into play, the impact of culture-specific knowledge is likely to be even more pronounced.

2B. COMPLEX ISSUES OF CONTENT VALIDITY

Content validity refers to the ability of a test to describe a person's performance on a prescribed universe of tasks. Problems of content validity arise when tests purport to assess one content domain (e.g., counting skills), but do so in a way that conflates performance in that domain with requisite performance in another domain (e.g., reading or verbal skills).

There are various test development strategies and statistical techniques for trialing, identifying, and eliminating those items that evidence significant test bias against any particular subgroup in the population cohort to be tested. Some efforts have been made to reduce the discrepancies in US racial-group performance through the use of different formats for items, including adjusted cultural content and non-verbal formats that attempt to retain comparable levels of cognitive, skill or task demand (e.g. Bracken & McCallum 1998). These reportedly have resulted in smaller differences of performance between cultural groups.

Generally, test bias refers to elements in a test that are not related to key constructs, and which appear to induce different levels or patterns of performance by group membership (Bracken & McCallum, 1998). For a test to be considered culture fair in conventional psychometric terms, it must meet standards of validity regarding both its internal technical characteristics and its relationship to the external characteristics of its sample, domains, etc. To date, considerable effort has been put into establishing

criteria for culture fair test construction. According to Bracken and McCallum (1998), contemporary strategies have included the development of:

- Tests that are less dependent upon verbal and linguistic knowledge;
- Tests which feature multiple rather than single indices of skill or ability;
- Tests with minimal emphasis on culture-specific background knowledge;
- Tests which de-emphasise speed of response; and
- Tests with varied format and response modes.

It is worth noting that there has been little explicit consideration of these matters in the development of currently used state-wide tests, despite an extensive Australian literature on problems of bias in the assessment of Aboriginal and Torres Strait Islander peoples. There were several early Australian attempts to develop culturally appropriate testing instruments and a considerable corpus of Queensland psychological research in the 1970s and 1980s to examine how thinking was affected by cultural knowledge (Boulton-Lewis, Neill & Halford, 1986). Researchers found that Indigenous Australian children performed less well on most tests of intelligence, although as contact with mainstream urban Australian culture increased, their performance moved closer to that of Anglo-Australian children (McIntyre, 1976). This early research also claimed that differences between Aboriginal and white children increase according to the age of the sample (Ross, 1984; Boulton-Lewis et al., 1986). Other studies found that Aboriginal children and adolescents performed better on some tests of spatial memory than did mainstream Australian children (Kearins, 1983). The Queensland Test (McElwain & Kearney, 1970, 1973) was an early attempt to address some issues of cultural specificity. Methods of obtaining a solution were not scored, test items were untimed, and the test was non-verbal, based on demonstration, unspoken instruction and mime. Further, scoring was based on three “contact norms” (Ross, 1984), which were designed to reflect the degree of western contact of the person being tested.

To conclude, there are no definitive approaches or answers in the search for culture-fair assessment. But there is a broad knowledge base on culture-fair assessment strategies. We know both how tests potentially discriminate against particular groups, and as well how they might be modified or altered to better accommodate cultural and linguistic diversity.

In both theory and practice, education systems have an ethical responsibility to ensure that assessment – regardless of whether it is high-stakes or not – is fair. What is obvious is that a good deal of research and development has been undertaken internationally to develop test development, administration and interpretation strategies that attempt to:

1. accurately describe the teaching and learning accomplishments and needs of all students; and
2. identify *individual* strengths and weaknesses for intervention, rather than reproduce larger subgroup differences.

We would argue that there are at least seven characteristic areas that require investigation if we are to move towards more culture-fair assessment in Queensland. These are, specifically:

- **Issues of language:** examining whether and how assessment in areas other than language and literacy (e.g., numeracy) are biased towards mainstream, Standard Australian English speakers;
- **Issues of cultural content:** examining how and whether assessment items and texts are biased towards the representation and portrayal of mainstream, non-Indigenous cultural practices and knowledges;
- **Issues of developmental sequence:** examining how the developmental sequences presupposed by continua and domain specification may be based on particular mainstream cultural patterns;
- **Issues of framing:** examining how the formats, presentation and administration of assessment items and tests may presuppose particular cultural knowledges and practices;
- **Issues of content and construct validity:** examining whether the particular skills and knowledges purported to be tested are indeed those assessed. In particular content validity refers to the adequacy of the coverage of the domain that an instrument purports to test (for example if a test fully identifies the scope of *literacy* within a given context). Whilst construct validity refers to the degree to which operationalisation of the instrument reproduces the theorised relationship between multiple indicators within a domain of interest (such as reading or dictation); and
- **Issues of interpretation and reporting:** examining how the population categories and groups used for disaggregating data may represent Indigenous performance in homogeneous and heterogeneous ways.

In what follows, we examine the contexts and background of current state-wide testing practices in Queensland schools with these issues in mind.

3: INDIGENOUS EDUCATION AND ASSESSMENT POLICY AND PRACTICE

3A. INDIGENOUS EDUCATION AND THE DEFICIT MENTALITY

The term 'compensatory education' emerged in the 1960s and 1970s to describe systems' responses to the apparent lacks, gaps and 'deficits' revealed in culturally different groups' performance on standardised measures. As we have argued here, one of the effects of psychological testing has been the translation of cultural difference into an educational focus on remediation and compensating for the educational 'deficit' of Indigenous children.

Having research based on psychometric testing as the basis for educational policy and programming for Indigenous Australians has led to a climate where:

- a gap between the achievements of Indigenous and non-Indigenous students is considered unextraordinary and normal by many in the general and educational communities (MCEETYA, 2000, p. 13);
- resultant programs are generally compensatory and remedial, aimed at changing the students to better fit the general system rather than being perceived as part of schools' core business;
- the models upon which educational programs are based are deficit-orientated in that they aim to locate difference as deficit and offer solutions of intervention to minimise this difference; and
- models of assessment are normative and prescriptive, assuming models of child development based on 'Western' children at best, and assuming that these models should provide a standard against which to judge Indigenous children.

Regardless of the particular intents of such approaches, the deficit mentality has had two significant effects on Aboriginal and Torres Strait Islander education:

- first, it has created an atmosphere in many educational settings where Indigenous students are expected to fail, when *all* of the literature on Indigenous and ethnic minority education internationally indicates that those classroom environments that have high expectations that all children can learn are most successful at turning performance (e.g., Ladson-Billings, 1994);
- second, it has marginalised Indigenous interventions and educators to 'add on', 'pull-out' or remedial status and deterred efforts to view Indigenous education and achievement as mainstream issues that need to be addressed via mainstream curriculum and pedagogic reform.

Many, far too many, Indigenous students are achieving below acceptable standards in a range of areas. Many, far too many, Indigenous students require specific

interventions and assistance. But this generalised dampening down of expectations resultant of deficit mentality, in part sustained by the reporting of standardised test results, has in fact deterred, rather than assisted schools and teachers from moving forward.

This is not to say that standardised testing and systematic diagnostic assessment in and of themselves are necessarily the problem. We know it is possible for instruments to assist in identifying children's learning needs and in providing appropriate interventions. And it is possible for us to use broad and large scale assessment data to track aspects of improvement of provision of education to Indigenous learners. But to better understand how standardised assessment has come to assume the focal role it presently enjoys, we need to examine its linkages with Indigenous education policy. In what follows, we provide a brief overview of Indigenous and Commonwealth and State education policy as a context for the rise of testing. This is followed by a review of the contexts and development of the Year 2 Diagnostic Net and finally a review of the contexts and development of the current state-wide accountability measures.

From Policy to Testing

In 1989, the National Aboriginal and Torres Strait Islander Education Policy (AEP) set out twenty one goals which aimed at achieving educational equality for Indigenous students by the year 2000 (Ministerial Council on Education, Employment Training and Youth Affairs (MCEETYA), 2000, p.5). While the AEP had as its standard that Indigenous students should have the same levels of access, participation and outcomes as the student population in general, these levels have not been achieved for this group in the decade following the policy's introduction.

In 1999 the Ministers of Education again agreed to National Goals for Schooling (MCEETYA 1999a) which continued the focus on literacy and numeracy basic standards, an outcomes based approach to schooling, and additionally highlighted the need for education to be socially just. Of particular interest to this review are the following goals:

- 3.3 Aboriginal and Torres Strait Islander students have equitable access to, and opportunities in, schooling so that their learning outcomes improve and in time match those of other students.
- 3.4 all students understand and acknowledge the value of Aboriginal and Torres Strait Islander cultures to Australian society and possess the knowledge, skills and understanding to contribute to, and benefit from, reconciliation between Indigenous and non-Indigenous Australians.
(MCEETYA, 1999a)

The Council also specifically noted the importance of all students – including Indigenous students – “having access to the high quality teaching necessary to enable completion of school education to Year 12 or its vocational equivalent and that provides clear and recognised pathways to employment and further education and training” (3.6).

To achieve these goals for Indigenous students, a taskforce was established which reported back to the Council in 2000, (MCCETYA, 2000) recommending, among other things, a model for more culturally inclusive and educationally effective schools. (MCEETYA, 2000, p.2) The continued focus of this taskforce, and the Council in general, on literacy and numeracy outcomes, has led to the implementation of the National Indigenous English Literacy and Numeracy Strategy 2000-2004 (Commonwealth Government of Australia, 2000). The objective of this strategy is to “achieve English literacy and numeracy for Indigenous students at levels comparable to those achieved by other young Australians” (p.5). The accountability elements of this strategy are closely tied to measuring educational outcomes in literacy and numeracy against agreed upon national benchmarks (p.32).

In 1999 a Queensland internal review of education and employment programs for Aboriginal and Torres Strait Islander Peoples reported that the current approach had been failing Indigenous Queenslanders. This led to the introduction of the *Partners for Success Strategy* (Education Queensland 2000a) and of six policies covering issues of standards, community involvement and partnerships, employment and whole of government service delivery. Of particular interest to this review is the Policy on Literacy for Aboriginal and Torres Strait Islander students which has as two of its objectives:

- to ensure that literacy teaching for Aboriginal and Torres Strait islander students is built on recognition of second and third language learning;
- to monitor and report student progress in ways that enable students, teachers and parents to collaborate in strategies for improving students’ learning.
(Education Queensland 2000a, p3)

There is, then, a clear move to increasing reliance on standardised outcome indicators of student performance as key measures of educational equity and inequality. It is noteworthy and interesting that since 1990, there has been no mention in either Commonwealth or State literature of the need for culture-fair or inclusive assessment strategies – despite the extensive international debate over culture-fair assessment noted in this report.

3B. THE NATIONAL MOVE TO OUTCOMES BASED EDUCATION

The Quality of Education Review Committee Report (Australia, 1985) argued that schools and systems needed to shift their emphasis from educational inputs to educational outcomes. The report also highlighted the need for a focus on ensuring that all students reached minimum standards at an early age in literacy and numeracy. This call for greater proficiency in literacy and numeracy and for a focus on the early years of schooling, continued as a policy direction throughout the early nineties (Department of Employment, Education and Training, 1991; House of Representatives Standing Committee on Employment, Education and Training, 1993).

Simultaneously, during the late 1980s the then Labor Commonwealth government made moves toward competency-based scales, particularly in relation to secondary

and post compulsory schooling. The creation of the National Training Board (NBT) and the establishment of the Finn Review of Post Compulsory Education advocated major developments to post compulsory education with a commitment to incorporate nationally agreed essential competency standards in all programs. The Mayer Committee was commissioned to further develop these curriculum competencies and in due time identified seven key competencies, claimed to be essential for effective participation in the modern work force. The tabling of this report (1992) can be contextualised within a political climate aimed at making the Australian economy more export orientated and internationally competitive. These policy decisions linked the need to restructure the workplace with the need for a nationally consistent, competency based education (Porter, Rizvi, Knight and Lingard, 1992, p.9).

These trends continue to be evident in Commonwealth government policy. Attempts to develop a national school curriculum framework for eight Key Learning Areas in the early and mid 1990s and the more recent development of National Literacy and Numeracy Benchmarks (Commonwealth Department of Education, Training and Youth Affairs 1997, MCEETYA 1997) are cases in point. The benchmarks and their development will be discussed in more detail later in this review, but we now turn to a discussion of what proved to be the catalyst for much of the curriculum reform levelled at Queensland schools in the 1990s.

3C. THE QUEENSLAND YEAR 2 DIAGNOSTIC NET AND INDIGENOUS LEARNERS

Curriculum Review and Reform

The Wiltshire Report's (Queensland Review of Queensland School Curriculum, 1994a) focus on early literacy and numeracy "was prompted by the recent history of concerns expressed by parents, educators and employers that some students lacked the levels of literacy and numeracy skills expected of them" (Queensland Review of Queensland School Curriculum, 1994b, p.2). Particular reference to literacy and numeracy appeared in one of the review's terms of reference:

TR 4: Identify the most effective forms of remedial intervention to help all children achieve their literacy and numeracy potential.

As expected the reporting of this review process included a clear focus on literacy, numeracy and the early age prevention, identification and remediation of failure in these areas. While the report made one hundred and five recommendations, those specifically related to literacy and numeracy were placed high on the timetable for implementation of the report's initiatives. Specifically the reviewers recommended the following of interest to this report:

- The introduction in all early childhood classes of teacher monitoring of student progress in literacy and numeracy through running records and the use of these records to enable teachers to seek formal, specialist diagnosis of those students who continue to experience difficulties. These records will also aid in

the assessment of resources needed to improve the literacy and numeracy of students;

- The introduction of a Year 2 early-age Net whereby diagnosis (suggested by the running records) of the literacy and numeracy levels of all students after 18 months in the compulsory school system will be made. For those students who have inadequate levels of literacy and numeracy, intervention in the form of intensive remediation to be introduced. The need to diagnose all children to be reviewed after three years;
- The introduction in Year 6 of a test for all students in literacy, numeracy and general skills. The test to be similar to that developed by ACER for other Australian systems and to include:
 - an individual report;
 - a report to parents on their child's test level;
 - a report to individual teachers so that they can diagnose those areas with which students are having difficulty.

(Queensland Review of Queensland School Curriculum, 1994a, pp.x-xiv)

The Year 2 Diagnostic Net

In reply to the recommendations of this review, the Year 2 Diagnostic Net was conceived as a means of diagnosing those students identified as having difficulty learning literacy and numeracy in the early years of school, and to provide remediation for those identified students. After some minor changes to organisational issues, the Net remains a four-step process in which teachers of children in Years 1, 2 and 3:

- monitor all children's progress using developmental continua for aspects of literacy and numeracy;
- validate their observations of Year 2 children using specifically designed assessment tasks;
- provide appropriate learning support for children requiring additional assistance;
- report to parents about their child's learning and development in aspects of literacy and numeracy.

(State of Queensland, 1996, p.1)

The whole process involves the teachers and students of Years 1 to 3, however the validation process and subsequent identification of students requiring additional support is specific to Year 2 only. The process is mandatory for all state schools, which relates closely to its political motivation (Grieshaber 1997) and its implementation in a climate of calls for teacher accountability in relation to standards in the basics of education.

After an announcement that the State government planned to implement the recommendations from *Shaping the Future* (Queensland Review of Queensland School Curriculum, 1994a) for early literacy and numeracy in February 1995, the State Cabinet set firm guidelines as to the conceptualisation and implementation of the early age net. The decision to use tasks instead of a test was seen as a positive

trade off in the teacher led movement against the introduction of a state-wide test such as that already being planned for Year 6, into the earlier years of schooling.

A decision was made that the *First Steps* materials from Western Australia would be used to satisfy the first component of the process for aspects of literacy. This chosen form of the Year 2 Diagnostic Net was certainly influenced by several considerations, however the tight timeline for implementation was definitely one of these considerations. Because no equivalent resource was readily available for numeracy, a Queensland team set about developing continua to be used for monitoring and validation in number.

The *First Steps* Program and Evaluations Previously Conducted

According to the developers' own literature (<http://www.ecurl.com.au/src/Firststep/main-firststep.html>) the *First Steps* materials are an English language resource developed in the late 1980s that "help individual teachers, schools and education systems achieve targeted literacy outcomes and standards for their students". There is a definite link made between assessment and teaching and learning, as the resources provide Developmental Continua that "link the monitoring, assessing and data gathering on student performance to the planning and delivery of the teaching and learning program"².

Earlier responses to such a perceived need to improve literacy standards and outcomes in Western Australia had involved whole language based teacher professional development in forms such as *ELIC* (Early Literacy Inservice Course). These approaches had been judged as being too dependent on the individual teacher initiatives, which had not necessarily received sufficient support at a whole school level. Nevertheless these previous attempts at renewal of the teaching of early literacy and numeracy were reflected in the model adopted for the *First Steps* materials. The materials draw heavily on the Vygotskian notion of zone of proximal development and Cambourne's conditions of learning (Cambourne 1988). This is evidenced in the problem solving approach to teaching and learning used, as well as notions of learning needing to be embedded, contextualised, scaffolded and based in interaction, with sufficient time provided for integration and consolidation.

First Steps has been criticised for relying on anecdotal evidence from teachers and schools to support its approach to teaching and learning. There has to date, been no independent review of student performance outcomes related to the project. No data on student performance or outcomes was collected before project implementation so possibilities for such review are now limited to comparisons between the results gained from schools who have been implementing the project for longer periods of

² An interesting point to note is that originally the Department of Education in Queensland only purchased the Developmental Continua i.e. the assessment tools of the package, and not the resource books. While teachers and schools could – and in many cases did – purchase them individually, this initial purchase decision highlights the value given to a decontextualised use of the materials as assessment tools.

time and like schools who are new to the project (Australian Council for Educational Research (ACER), 1993). During the period of 1992 – 1994 the Australian Council of Educational Research did review the project and its materials with a focus on the technical validity of the Developmental Continua and on the survey of participants' opinions about the project's implementation.

In 1994, three further research projects were commissioned to a research consultant. They entailed: the documentation of the history of *First Steps* development up until 1994; a survey of 150 schools to ascertain the opinions of school based participants in the project in relation to its level of implementation and the effect of this at local schools; and the development of twelve case studies to investigate the effect of the project across school locations. While the Executive Summaries of these studies recognise their limitations, in general they cite positive results for the *First Steps* project (Deschamp 1994 and 1995).

Of particular interest to this review is a report of the Implementation of the Literacy Component of the *First Steps* Project in ELAN Schools (i.e. schools with high numbers of Aboriginal students). The Executive Summary of this review (Deschamp 1994) clearly states that teachers' responses ranged from very positive to very negative when asked about the project's effectiveness in relation to assessing and teaching Aboriginal students. However the summary of results listed as the final section of the summary provide the following positive findings of the review:

- Teachers in ELAN schools were even more positive about the value of all aspects on the *First Steps* professional development than teachers generally;
- Teachers in ELAN schools valued all aspects of the *First Steps* model even more highly than teachers generally;
- Teachers in ELAN schools gave higher ratings than teachers generally to all aspects of the impact of *First Steps* on their teaching. This suggests that they considered *First Steps* has particular relevance to teaching children who needed extra assistance with learning English;
- ELAN schools had made similar or greater progress than schools generally with regard to the extent to which *First Steps* literacy program has been implemented. The greatest differences were in Writing and Reading, where ELAN schools claimed substantially more progress with implementation than schools generally;
- Principals in ELAN schools gave higher ratings than principals generally to the impact of each aspect of the *First Steps* literacy program on student learning.

Like the *First Steps* research in general, the above review was not based on any systematic meta-analysis of student performance, but rather on the views of teachers and administrators.

In 1995 the Northern Territory Department of Education also commenced a program of implementation and formative research on the *First Steps* materials. Because of the Northern Territory's large Aboriginal student population, this research is important for our present purposes. The research aimed at investigating the impact of the complete program on students and teaching, and as such focused on the teaching strategies, resources and the whole school planning supported through the program, rather than

on the cultural appropriateness of the Developmental Continua as such. In fact no analysis of the appropriateness of using the continua with specific groups of students was undertaken as part of the external evaluation.

The original external evaluation involved three projects. The first was a series of twelve school case studies conducted over a period of three years. The purpose of this research was "to document how the schools were teaching literacy before they commenced using *First Steps*" and to compare this to how the teaching of literacy was conducted after the teachers had been trained in using *First Steps* materials (Northern Territory Department of Education 1999, p.vi) The report of this research detailed what was happening in schools where the levels of implementation of the materials varied, and provided probable reasons for this differential coverage. The recommendations of the report are more related to developing whole school strategic plans and having whole school support of these plans than specifically about the benefits of *First Steps*.

The second Northern Territory research project involved "collecting hard data about students' test results" before school-based implementation of *First Steps* and comparing these to student results after the school had been implementing the program for some time (Northern Territory Department of Education, 1999, p.vi). The hard data was collected by using standardised tests - the *Waddington Reading Test*, the *Student Outcome Writing Test*, and the *Easy-Mark Diagnostic Spelling Test* - before and after *First Steps* implementation at selected schools.

The results demonstrated a slight increase in students' outcomes in spelling and large gains in students' outcomes in writing. *Students' outcomes in reading actually declined*. The report claims that the results obtained in this project are "the most powerful support for *First Steps* to date" (Northern Territory Department of Education, 1999, p.5) due primarily to the unique position of the Territory to monitor the impact of the implementation through the collection of baseline data. The decline in students' outcomes in reading as measured by the *Waddington Reading Test*, is explained as the result of the fact that the writing strand of the project had been first implemented in the Territory, with the reading strand only very recently implemented during the post-tests period.

The third research project involved a survey of Principals, teachers and parents within a sample of schools to determine the extent to which *First Steps* was operating in the schools and what they considered had happened as a result of this implementation. This research project mirrored those already conducted in Western Australia relying on opinions and reflections - and thus did not investigate issues related to the present study.

In addition to the reports from the research detailed above, the Northern Territory Department of Education has published an edition which documents case studies and testimonials of more than twenty contexts where the *First Steps* program is used across the Territory (Northern Territory Department of Education, 2000a). While responses to the program vary, the report is generally positive of the impact of the *First Steps* implementation. It is interesting that there is very little mention of the use of the *First Steps* materials with Indigenous students or the appropriateness of this.

One comment made is from an ESL teacher who teaches at a large Alice Springs school, which caters to a 25% Indigenous population:

From an ESL point of view, there has always been a rift as to what extent First Steps' Developmental Continua can be used to track second language development.

(Northern Territory Department of Education, 2000a, p.57)

The Northern Territory Department of Education also implemented a *First Steps* Intervention Project during 1998 and 1999. This study investigated the use of *First Steps* strategies for individual children in pull out or within class extra assistance programs in urban, small rural and remote school contexts. The research reported generally positive results in student outcomes and in teaching and school based development (Northern Territory Department of Education, 2000b, pp.8-11). Of interest to this review is the fact that while the use of the *First Steps* strategies were reflected as positive for all the students involved, *the Developmental Continua were not considered by teachers as appropriate tools to reflect the progress of Indigenous learners*, with teachers choosing to use the ESL Outcomes Profiles for this group instead (Northern Territory Department of Education, 2000b, p.8).

The *First Steps* materials are currently being rewritten, and updated, with a focus on better addressing the learning needs of Aboriginal students, students for whom English is an additional language, students with learning difficulties and students living within rural and remote areas (Footnotes, 2000). *To date systematic research has not been carried out investigating the appropriateness of the First Steps project for Indigenous students.*

Evaluation of the Year 2 Diagnostic Net

The Queensland Year 2 Diagnostic Literacy Net is based on the *First Steps* Developmental Continua. As previously mentioned the Net process has been adapted throughout its years of implementation. In 1995 all students were validated as part of the Year 2 Diagnostic Net validation process, however in 1996 a decision was made to limit the teacher work load involved in the validation process by no longer requiring that all students be validated. Instead the validation process was only required for those students who were expected to be identified as requiring extra support. Changes were also made to the specific criteria used for identification of students requiring extra assessment after the 1996 validation process. However, there have been no systematic adaptations to date for Indigenous learners.

In 1997 the first of two reviews of the Year 2 Diagnostic Net process was commissioned by the Queensland School Curriculum Council (QSCC). A consultancy team from the University of Queensland was directed to provide an intrinsic critical appraisal of the Year 2 Net (Luke, Land, van Kraayenoord and Elkins, 1997). The report of this appraisal was delivered but not released. The original report noted questions about whether the developmental continua were narrow and monocultural, and whether they accommodated the diverse pathways of learners of other cultural and linguistic backgrounds. In 1999 the researchers published a brief critical commentary on the Year 2 Diagnostic Net. They documented concerns with the

process in relation to the basic assumptions made by the materials and processes in relation to: the view of learners and literacy learning; assessment; and categorisation of at risk categories. They also documented the problematic nature of the relationship between the Year 2 Diagnostic Net and the current Queensland English Syllabus, issues of validity and the moderation process, and concerns about the potential uses of data collected through the process (van Kraayenoord, Luke, Elkins and Land 1999).

In 1998 a second review of the Year 2 Diagnostic Net was commissioned to a consultancy team from Central Queensland University. The design of this review focussed on the:

- appropriateness of use of 'required resources' in meeting clients' needs especially those of teachers;
- effectiveness of use of material resources to meet the purposes, processes and outcomes of the Year 2 Diagnostic Net; and
- efficiency of teachers' use of 'required resources' to achieve stated outcomes.

(Stewart-Dore, Bartlett, Hallinan, Moreton, Robert and Woodrow, 1999, p.iii)

This review team noted that the minimal level of resources available to conduct the review, and the design of the research, allowed only open and broad interpretations from participants. This they said limited their ability to conduct "close analyses of specific issues such as reporting" (p. iv).

Both reviews raised issues about the appropriateness of the Year 2 Diagnostic Net materials for Aboriginal and Torres Strait Islander students. The CQU summary specifically notes the linguistic differences between the uses of varieties of English by these students for whom English is a second, third or fourth language and the materials' focus on Standard Australian English (SAE).

The Need to Review Net Materials and the Developmental Continuum

We have documented that there have been questions raised about the effectiveness and appropriateness of the *First Steps* Continuum and the Year 2 Diagnostic Net Process in separate, unrelated studies undertaken in Western Australia, Northern Territory and Queensland. These studies have not been definitive – but they have noted concerns that reflect the categories of potential cultural assessment bias noted earlier in this report. These categories include:

- **Issues of Language** in relation to whether the Net materials and assessment format have differential effects on Creole, vernacular and second dialect speakers;
- **Issues of Cultural Content** in relation to whether the content of the actual text materials are culture-specific; and

- **Issues about Developmental Sequence** and identification of the fact that because the Net actually maps children in a criteria based format in relation to a sequence of skills, competencies, knowledges and behaviours, it might not accommodate or might misclassify alternative cultural pathways to literacy, which now have been extensively described in the literature (e.g., Gregory, 1997).

How does the current Year 2 Diagnostic Net administration accommodate Aboriginal and Torres Strait Islander learners? The information provided to schools to direct their conduct of the validation process does contain a short chapter relating to socially just practices (Department of Education 1995, pp.26-37). Within this section the learning needs of Aboriginal and Torres Strait Islander children are discussed with particular note being made of the following issues:

- that teachers using the developmental continua for ATSI children should be supported with cross-cultural awareness in-service to make them aware of appropriate methods of interacting with ATSI children, thus ensuring that these children's development is accurately mapped on the continua;
- that the contextualisation of the validation text take into account the perspectives, contributions and experiences of all children, and that the activities selected be inclusive and culturally appropriate to all children;
- that the teacher establish a warm and positive relationship with the child;
- that the teacher be cognizant of the fact that ATSI children may speak an Indigenous language as their first, second, third or fourth language, and ensure that the child understands the requirements of the task;
- that teachers remember the fact that an ATSI child's completion of the validation task may reflect features of non-standard English. It is suggested that such variations may be considered as indicating the need for further support in learning (Standard Australian) English rather than indicating a learning difficulty;
- that hearing loss and other health problems may affect a child's performance on the validation tasks; and
- that the teacher consider eye contact and direct questioning when conducting the validation tasks with ATSI children - as these can be "culturally inappropriate and threatening".

(pp.28-29)

These may be sensible recommendations about the assessment administration and validation procedures. But there have been no dedicated training sessions for teachers with Aboriginal and Torres Strait Islander children to date. And the actual materials and developmental continua remain unchanged, despite at least two calls for further consideration of this issue.

To summarise then, it would seem that there has been little rigorous empirical analysis of the performance patterns of Indigenous students as measured by the Year 2 Diagnostic Net process. Both commissioned reviews (Luke, Land, van Kraayenoord and Elkins 1997; Stewart-Dore, Bartlett, Hallinan, Moreton, Robert and Woodrow, 1999) of the process had limitations placed on the availability of data. Both, like other

WA and NT reports, noted the need for a specific review of relevance and appropriateness of both the First Step's Developmental Continua, and more generally the use of singular pathway continua for assessment of Indigenous learners. *Such a review is now well overdue, and remains crucial as the assessment processes for early years are melded with curriculum in the new syllabus documents.*

3D. THE QUEENSLAND YEAR 3, 5 AND 7 TESTING PROGRAM AND INDIGENOUS LEARNERS

Literacy and Numeracy Benchmarks

In 1997 all State and Commonwealth Ministers of Education agreed to sign a new National Goal for education. It stated that:

- every child leaving primary school should be numerate, and be able to read, write and spell at an appropriate level.

They also agreed that:

- every child commencing school from 1998 will achieve a minimum acceptable literacy and numeracy standard within four years.

(MCEETYA 1999a)

This led to the endorsement of the National Literacy and Numeracy Plan, which aimed at co-ordinating the approach taken to improving literacy and numeracy standards across States and the Commonwealth. These measures in turn led to the assessment of students Australia wide, using rigorous but state-based assessment procedures.

The benchmarks are part of these agreed moves to improve the educational outcomes of all Australian children. They are a set of indicators or descriptors which represent nationally agreed minimum acceptable standards for literacy and numeracy at a particular year level (<http://detya.gov.au/schools/literacy&numeracy/benchmarks.htm> last updated July 2001).

Benchmarks do not claim to represent the full range of the curriculum, but instead only the essential elements of literacy and numeracy. This claim of essentialness is problematic (Christie, 1998). That these essential elements of literacy and numeracy might be context dependent is a key issue. Much of the current curriculum innovation within Australian education systems is based on the recognition that students need an understanding of, and ability to use appropriate skills in context, and yet the benchmarks seem to be based on the assumption that a narrow set of universal skills regardless of context will prepare students for their future lives (Luke and van Kraayenoord, 1998, p.60). Similarly, there seems to be a level of acceptance that if it is possible to document these essential elements, that the descriptions in the official Curriculum Corporation's publications are such documentation. With no notion of who it is being assumed these benchmarks are essential for, or for what purpose, and

on whose authority, the benchmarks in their present form depict a monocultural representation of literacy and numeracy competence (Christie, 1998, p. 44).

The benchmarks are based on assumptions of accountability and improved outcomes. Firstly it is assumed that the articulation of minimum national standards will lead to increased system and school accountability to key stakeholders. Secondly it is assumed that the translation of these minimum standards into a rigorous state-based assessment process will lead to improved outcomes for student achievement and improvements in teaching quality (Luke and van Kraayenoord, 1998). There is to date little evidence that system based testing regimes actually have the capacity to lead to improvements in these areas. Willis (1998) states that rather than enhancing student outcomes, the benchmarks in numeracy are more likely to undermine improvements to student outcomes. Generally the benchmarks are not based on an adequate conceptualisation of what factors might put children at risk of not learning, and because of their minimum competency approach to performance description and narrow back to the basics focus, they may well be more likely to “undermine good teaching practice than enhance it” (Willis, 1998, p.71).

Secondly, the very nature of the performance testing and the necessity to test cohorts of large numbers of students will always constrain test development and the development of essential criteria used to define complex concepts. Paper and pencil tests are limited in what and how they are able to assess.

Queensland's Approach to State-wide Testing

Beginning in 1999, the results of these testing processes have been reported annually within the National Report on Schooling in Australia (MCEETYA, 1999b). Queensland state schools now have a State-wide testing procedure mandated for all children in Years 3, 5 and 7. This testing program is aimed at accounting for and contributing to the improvement of student learning in aspects of literacy and numeracy (Queensland School Curriculum Council, 2001, p.1). The aspects of literacy tests now cover four strands – reading, writing, spelling and viewing, and were developed in accordance with the Queensland English Syllabus and the national benchmark statements (Queensland School Curriculum Council 2000). The aspects of numeracy tests cover three strands – number, measurement and data, and space, and again were developed in accordance to syllabus documents and national benchmark statements.

The tests are standardised, with response modes covering multiple choice, short answer and long answer formats. The majority of the tests are machine scored, however some sections such as the writing task are marked by individuals bringing issues of moderation and inter-rater reliability to the foreground. The census tests are criteria referenced to syllabus objectives, although not exhaustively or sequentially. Student performance is norm referenced.

While the tests are defined as ‘low stakes’ at an individual level because students' access to educational opportunity is not necessarily influenced by performance on the tests, this is not the case at school level. Within the state system, student performance must be reported within each schools' system based reporting processes. In this regard

the tests are high stakes at least for schools, with student performance on the tests being used to negotiate funding and also being used to make judgments about the efficacy of administration and teaching staff. While the QSCC claims that the tests are designed to inform teaching decisions, the timing of test administration and results are not conducive to their use for diagnostic and formative assessment purposes.

Analyses of state-wide mean literacy and numeracy performances for key student groups allows comparisons between groups of students on the basis of gender, Indigenous background, language background, and living environment (urban or rural). In 2001 the mean scores of Indigenous students on all strands of both literacy and numeracy were lower than those of non-Indigenous, in all three year levels (Queensland School Curriculum Council, 2001). The extent of these differences was such that the mean results of Year 7 Indigenous students on both literacy and numeracy were akin to the mean results of non-Indigenous students in Year 5. In the same way the mean results of Year 5 Indigenous students were similar to those of Year 3 non-Indigenous students. Interestingly the QSCC interprets these results to suggest that school authorities continue to be concerned about these large mean differences, and that they continue to devote special attention to this group (p.38). No mention of the effect of the test items and testing procedures on the performance of Aboriginal and Torres Strait Islander students is made in the annual reports.

While the QSCC might not make mention of the appropriateness, or otherwise, of the use of the Years 3, 5 and 7 tests with students of Indigenous background in their reporting procedures, teachers and school administrators did make mention of the same when asked to comment on the 1998 Years 3 and 5 testing materials (Queensland School Curriculum Council 1999). In a survey conducted by the QSCC, teachers and administrators reported that the Year 3 and 5 test materials were not inclusive in several ways. Participants made comment about the inappropriateness of the content, language, and testing style and format for Indigenous students. It is unclear whether the responses of these school-based personnel have been considered during test development in following years.

In what follows, we describe a systematic investigation into the 2000 Queensland Year 5 test, examining issues of linguistic and cultural content, construct validity and the analysis and reporting of results from the perspectives on culture-fair testing raised earlier in this report.

4: 2000 QUEENSLAND YEAR 5 TESTS IN ASPECTS OF LITERACY AND NUMERACY - A CASE STUDY

As noted above Queensland currently runs comprehensive literacy and numeracy testing programs in compliance with its benchmarking agreements with other States, Territories and the Commonwealth. This study requested detailed item-by-item data on the 2000 results on all tests for statistical and content meta-analysis. Education Queensland provided us with detailed data on the Year 5 Aspects of Literacy and Numeracy tests only. In consequence, we opted to do a fine-grained analysis of this test.

The design of the following case was meant to combine the best of Aboriginal and Torres Strait Islander epistemological viewpoint and professional expertise with state-of-the-art statistical/technical analysis. The purposes of this analysis were to:

1. test the criteria and framework we developed in the initial section of this report for assessing culture-fairness of assessment instruments; and
2. actually engage Indigenous educational expertise, complemented with high level statistical/technical expertise in a combined content/language/validity analysis. The latter exercise, it was hoped, would provide us with a sense of the viability of prototypical research and development procedures for the development of the 'next generation' of Queensland state-wide tests.

4A. INDIGENOUS EDUCATORS' PANEL

A panel of 7 expert Queensland Indigenous educators of various trainings, backgrounds and systemic responsibilities was convened to meet with Luke, Land and Woods for two, day-long working seminars in Brisbane. The research team had already made several field visits to schools to discuss the use of the Diagnostic Net, tests and appropriate reporting strategies. Recounts of issues raised in these field visits and the literature review provided input and catalysts for discussion of issues around assessment and testing.

The seminars consisted of open-ended discussions of language and literacy education, Indigenous education policy issues, and detailed analysis of the Year 5 literacy and numeracy tests. The panel consisted of:

- Mr. Lloyd Appo, Teacher, Bundaberg West State School
- Mr. Will Davis, Teacher/Coordinator, Murri Matters Program, Beenleigh State High School
- Mr. Steve Foster, Principal, Badu Island State School
- Ms. Maureen Liddy, Principal, Hopevale State School

- Professor Martin Nakata, University of South Australia
- Mr. Chris Sara, Principal, Cherbourg State School
- Ms. Grace Sara, Teacher, Cherbourg State School

The procedures were as follows:

- The expert team was presented with overall state test data and discussed their general impressions about the uses of the state-wide tests, their administration, and the reporting of results;
- The expert team was given copies of all test materials for the Year 5 Aspects of Literacy and Numeracy tests;
- The team was asked to score each item on a 1-3 rating, with 1 assigned to cultural content/language neutral items and 3 assigned to those items that they considered to be potentially biased against Aboriginal and Torres Strait Islander children's knowledge and competence;
- Each item was discussed in terms of fairness and bias and composite ratings were compiled;
- Exemplary items that the panel viewed as particularly problematic were singled out by the research team for: (a) linguistic and content analysis; and (b) statistical meta-analysis.

The Indigenous Panel's View of the 2000 Year 5 Aspects of Literacy and Numeracy Tests

The panel rated approximately 75% of all items on the literacy and numeracy tests as potentially discriminatory against Aboriginal and Torres Strait Islander students, particularly those who have vernacular, Creole and non-standard dialect as their first language, and those who live in remote communities. Their general sense of the tests was that they relied very much upon urban experience and were linguistically contingent at all levels.

The group expressed strong consensus on the following points, which are here categorised in terms of the framework for assessing culture-fairness introduced in Section 2B:

1. **Issues of language:** Several of the panel were Torres Strait Creole and Aboriginal English speakers. They argued strongly that even though students might be proficient at aspects of literacy and numeracy, specific linguistic features of Creole and vernacular (different case, plural, syntactic/tense structures) meant that the EFL/ESL/ESL speakers would have difficulty with both: (a) specific idioms, vocabulary, and signs that were urban and mainstream Australian; and (b) the case, conditionals, modals and transitivity structures of the test instructions and item 'distractors' within the tests. Over 75% of items were seen as potentially problematic for bi and trilingual Indigenous students.
2. **Issues of cultural content:** The panel noted that the actual cultural content of the test items very much relied upon student background knowledge of

specific experiences, lifestyles and practices common of mainstream urban Australian life. This included, for example, references to aspects of urban transit, pet ownership, urban institutions and leisure activities. It also noted that many of the numeracy tasks did not accommodate typical local ways of measuring, counting, assessing and evaluating data at which many Aboriginal and Torres Strait Islander children are extremely competent. Over half of the items assessed were ranked high in terms of potential bias on the basis of culture and urban-specific cultural content.

3. **Issues of developmental sequence:** The panel noted that the assumed developmental sequence of literacy and numeracy development underpinning the tests might disadvantage particular groups of Indigenous students in particular and specific ways.
4. **Issues of framing:** The panel noted three areas of potential difficulty in terms of test layout and framing. These were: (a) the complexity and contingency of the verbal test administration instructions for ESL/ESD speakers; (b) the difficulty on the literacy test of shunting back and forth between two texts (the mock-magazine and the test administration booklet); (c) that the test items seemed to vary in linguistic, semiotic layout and cognitive load format almost on an item-by-item basis. These issues meant that students were unable to develop clear schemata for the specific genres of a task or problem before they would be required to move on to another. In other words, it would be difficult, the panel believed, for Indigenous children to ‘get into a groove’ of a test genre, because of the complexity and multiple item format of the tests. The situation, the panel believed, was in part the effect of the test trying to cover so many curricular outcome domains.
5. **Issues of content and construct validity:** The panel noted that the numeracy tests were complex tests of reading comprehension and English-language fluency; and that some of the cognitive and behavioural demands of both literacy and numeracy tests did not match the construct and domain specifications that specific items set out to assess.
6. **Issues of interpretation and reporting:** The panel noted that the composite reporting of 'Indigenous' versus non-Indigenous achievement on a State-wide basis was not helpful in documenting where and how gains had been made. Further, while more helpful, there was some doubt that the ‘Like-Schools’ comparisons were as accurate as some would maintain. They argued that other ways of disaggregating Indigenous scores (e.g., by distance from urban centre, by ESL/ESL/EFL population) should be explored and that a dialogue around comparative test data, the setting of targets and implications for funding needed to be had – rather than simple judgements about program and school efficacy made solely on the basis of single-shot assessment.

To summarise then, generally speaking, the panel was not opposed to standardised testing of Indigenous children. Panel members felt that the Year 2 Diagnostic Net, if culturally modified, was a valuable instrument for teachers and students to engage with as part of rethinking their own strategies. The panel said that such instrumentation was useful for them as educators, for their communities and for

parents in assessing where they stood in relation to others, and in setting targets and goals for their schools. But they felt that the tests were, from both Indigenous and non-Indigenous perspectives, not technically of high-quality.

It was felt that the current tests – as exemplified by the 2000 Year 5 tests – were too complex, and technically problematic in several areas. The panel argued strongly that some system-level statistical and content analysis of test-bias needed to be undertaken on a regular and much more systematic basis than occurred at present (This process is currently undertaken through periodic convening of ‘reference groups’), and that the State could, with a small investment, prototype and trial some alternative items that might be used to see if and in what ways Indigenous children fare better, if cultural and linguistic content is modified, while maintaining construct validity and criterion referenced standards equivalent to current test items and domains.

Finally, the Panel suggested that it would be very interesting to trial items for mainstream non-Indigenous children based on Aboriginal and Torres Strait Islander cultural content and language, to see if any culture-fair test bias issues were evident in the outcomes via large scale statistical analysis.

4B. A LINGUISTIC PERSPECTIVE

Language issues

As noted above, the Panel rated and discussed those items that it believed might induce test bias on the basis of linguistic diversity of students. The research team ‘banked’ those items that were deemed problematic on linguistic basis, and identified three broad categories of linguistic contingency that might have an impact on test performance. These are:

- Complex literacy load within items with domains not related to literacy;
- Overloading of clause complexity in item instructions and task framing, e.g., through the addition of modal auxiliaries and complicated negatives;
- The assumption that all candidates sitting the tests speak Standard Australian English as their first language to a level of near-native proficiency.

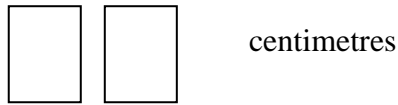
We then selected from the bank, specific items for closer linguistic and content analysis. In what follows, we provide exemplary item analysis to illustrate these problems.

Complex literacy load within items with domains not related to literacy

Numeracy - Item 14

Item 14 reports its construct as: the ability of students to *use a measurement mat to calculate the perimeter of a shape*. Yet a close examination of the item indicates the substantive contextual reading comprehension involved in the presentation format that is pre-requisite to being able to address the numeracy task.

14. Chris wants to decorate the edge of the letter L
Find the PERIMETER of the letter L



Taken as a series of three items (items 12-14), the embedding and contextualising of the measurement questions in supposed real life reported activities becomes obvious as an overt attempt to follow constructivist principles. However, this contextualisation raises the literacy load of the question, potentially disadvantaging those students having difficulty in literacy in Standard Australian English.

It also alters the construct of the numeracy task, converting an assessment of skill at measurement into an assessment of a student's understanding of particular forms of life experiences. Our point is that students who are struggling either with basic decoding or with reading comprehension - and the Years 3, 5 and 7 test data itself would suggest that this is disproportionately Indigenous students - would have a greater degree of difficulty demonstrating numeracy proficiency on such test items. This difficulty would increase proportionate to the reading comprehension demands of each item.

Overloading of clause complexity through, for example the addition of modal auxiliaries and complicated negative constructions

Modality of the text

Literacy - Items 28 – 32

Items 28 through to 32 ostensibly set out to assess students' abilities and skills in the viewing aspects of literacy. The items relate to an illustration from the front cover of a relatively age appropriate narrative text. There are issues in relation to the contextual appropriateness of the illustration – particularly for students living in remote areas not serviced for rubbish disposal by large garbage trucks. However the more general concern is that while the items are designed to assess viewing, the written presentation mode is also implicated in the performance of students. In this set of items the modality of the written items also ensures an increase in complexity in the literacy load required to achieve a correct response.

29. Which clue suggests the boy could be the main character?

The modality within this question is weak resulting from the use of the modal verb *suggests* and the modal auxiliary *could*. Besides increasing the very number of words required to be read to answer the question, the weak modality also decreases the certainty of the correctness of the available multiple choice responses. Similar concerns can be identified in several other questions in this set, where the use of modal auxiliaries and modal verbs suggest shades and levels of certainty and doubt. Because this particular construction of modality is less common in Torres Strait Creole and Aboriginal English, we would speculate that it could act as a distractor that will skew results by Indigenous ESL/ESD status.

Complicated clause constructions
Literacy – Item 35

Item 35 claims to be assessing students' ability to understand the effect of a word in context, however because of a difficult sentence construction, resulting from the linking of clauses by the use of an adversial conjunction (but not) much of the literacy load involved in the question is situated in reading the question rather than in understanding any word in the context of the reading text *The Paralympic Games*.

35. Which of the following sports **will be** part of the year 2000 Paralympics **but not** the year 2000 Olympics? (bold not in original)

Again, it was the Panel's belief that Torres Strait Creole and dialect speakers would struggle with complicated clause constructions such as this one, which, for all students, would not be common or in everyday use.

Language issues resulting from the assumption that all candidates sitting the tests speak Standard Australian English competently

The panel identified various words or phrases (particularly but not exclusive idiomatic uses) as likely to cause difficulty to students who do not speak Standard Australian English as their first language or dialect. As an example in item 3 the term *couple of weeks ago* would be a confusing use of language for speakers of Torres Strait Creole. Similarly the proofreading section includes several items in which confusions could occur because of the phonetic elements of vernacular spelling patterns in Torres Strait Creole, (as examples Proofreading Item 12 – neet for neat and Item 16 – mached for matched).

Within the numeracy test similar problematics result from the use of comparatives such as *taller than* and *shorter than* in measurement items (Numeracy Items 19). Members from the Indigenous Educators' Panel identified that within the language use patterns of Torres Strait Creole, the concept of tallness was represented with the word big, and not through the use of such comparatives.

Content

The Panel found that the 2000 Year 5 Aspects of Literacy test, by the very contextual nature of reading and writing practices, posed more cultural content issues than its numeracy counterpart. When dealing with issues of content it is important to hazard against token changes to content, which may in fact not noticeably increase the culture fairness of the assessment measure. The aim would be not to include token Indigenous or multicultural content (as has been undertaken in textbook and curricular revision). *In this regard, neither the expert panel or research team was concerned with what could be construed or termed 'political correctness' of test item content.*

Rather in the interests of better representing Indigenous students' performance, the argument made was to expand and improve tests to include content which is inclusive of students from backgrounds other than those of urban middle class. Within the 2000

Year 5 Aspects of Literacy Test reading material, it is possible to find numerous concepts that would be foreign to the experiences of children in rural or isolated communities. As examples - visiting National Parks as a *weekend away*, the RSPCA, public transport particularly ferries and City Cats, and Garbage Disposal Services, all appear as key concepts in the reading material. Simply put, the strong focus on these concepts and knowledges violates one of the first principles of culture-fair testing noted in our introductory section. That is the framing of test items in exclusive or extremely culture-specific content. *From cognitive approaches to learning through to theorising reading comprehension, we know that content counts - content makes a difference in text accessibility* (Spiro, Bertrum and Brewer 1980, Bruner and Haste, 1987)

Literacy - Items 1-6 and 13-16

When investigating issues of content appropriateness it is not however sufficient to view each text as a separate entity. A brief analysis of the content area covered in the texts positioned on the first double page of the reading material for the 2000 Year 5 Reading and Viewing test, highlights the confusing nature of the texts when taken as a collective. The first story *Relax Max* has, as one of the prominent characters, a dog called Cat, which would seem to be unnecessarily confusing to students when the domains of testing relate to making links between sentences and locating information in text. However when linked to the other texts on the adjacent page of the reading material the content based problematics become even more complicated. The visual texts under *Advertisements* include a small dog labelled as an unrestrained killer, a puppy who readers are requested to not classify as a toy, and a dog - who readers are expected to infer is named Rover - and who is being used in an advertisement related to City Cats. The confusing combination of these concepts would be expected to increase the literacy load of these texts, and the questions that relate to them, beyond what is necessary, and certainly beyond what the item domains purport to be assessing.

Literacy - Items 17 - 21

A more general concern would be the actual selection of texts. The text *Fancy 'Fido the Quoll' for a pet?* raises issues of content in relation to the uncommon subject matter contained in the text. Members of the Panel highlighted the fact that a passage on a more common animal may well have been a more appropriate context in which to assess the item domains listed. In addition to this however the format of the text contains several superfluous elements, including text boxes containing text which seem to be masquerading as hyperlinks. The notion of presenting reading material across modes and including multi media elements - which would seem undoubtedly to be the objective behind the inclusion of texts such as *Fancy 'Fido the Quoll' for a pet?*- may have seemed to be justified by current literacy theory. However the multimedia elements are actually not transferred across to a singular written format and thus become nuisance rather than addition to the text.

Framing and Context: Test Administration Issues

Schools are guided in their approach to administration of the Queensland Years 3, 5 and 7 Tests in Aspects of Literacy and Numeracy by information provided in *Information for Schools* (QSCC 2000). While this publication seems to set out the guidelines for administration in a standardised fashion, reports from the field gathered from the case study schools, supported by information collected during the panel and discussions with the QSCC's testing personnel, has led the researchers to make the contention that uneven deployment of these guidelines at a school level raises issues in relation to equality. A further contention would then be that it should become the responsibility of the QSCC (or its equivalent) to supply guidelines that are simple to decipher, less ambiguous, and less locally malleable.

To illustrate this point, a closer examination of the guidelines given for the issuing of *special considerations* for accommodation of the test conditions will be investigated. Schools can - and do - interpret the issue of accommodating special requirements for particular students according to two quite different criteria, depending on which section of the guidelines they choose to comply with. Schools are provided with details of categories of students who might require special consideration in test administration with minimum criteria for this categorisation and suggested special considerations that might be appropriate for students thus categorised (QSCC 2000, pp. 47 - 48). When adhered to these details ensure a limited access to special considerations for test administration. For students of Indigenous backgrounds the minimum criteria would certainly deny some students access to special considerations for test administration as many, particularly urban Indigenous students, may not meet the criteria as they presently stand.

For example many urban Indigenous students may not have an *ESL report on English language proficiency* (QSCC 2000, p.48) written on their English language skills even if Standard Australian English is not in fact their first language. Given the seasonal migration and family relocation of many Indigenous students, there is evidence that the EFL/ESL/ESD language map of the state actually shifts continually (Luke, Land, Christie & Kolatsis, 2002), and that many schools are not able to name, identify or classify Indigenous linguistic diversity. This issue of language diversity – given the item specific issues about linguistic load, construct validity and cultural content raised above – bears serious study and consideration in the formulation of test administration guidelines for both special consideration, exclusion and administration.

However if the school should instead decide to comply with the guidelines as described under the sub heading 'Use of support staff', access to special considerations is far more readily available as described in the following statement:

In general, individual students who are usually assisted in some way by support staff in the classroom may have their normal level of support in the test situation (if this is appropriate).

(QSCC, 2000, p.45)

During discussions with the research team, the QSCC's testing personnel acknowledged the ambiguity of this manual, and indicated that it had become problematic that interpretation of the information at a school level varied across the

state. At the same time, the field case studies here indicated that test administration conditions and contexts are highly variable. The accounts from the field range from one-on-one help (trained and untrained) in some community schools, no support in others, and the administration of the tests without any special consideration in some community schools which, reportedly, led to chaotic patterns of participation and overall diminished performance.

There is no simple answer to this issue, other than that clear guidelines, clear, empirically based criteria (e.g., language, indigeneity) for the identification of exceptional cases, and the 'norming' of special consideration intervention needs to be achieved. This will require some effort by test developers and analysts and administering bodies. It is worth noting that in the 2000 national testing reported by DEST (DEST 2002), Queensland had lower rates of exclusion from testing among Indigenous children than other states with significant Aboriginal and Torres Strait Islander populations. Without further detailed research and development on framing and test administration, we cannot know whether there has been a skewing of Queensland Indigenous results by the exercise of differential exclusion and special consideration categories on an intra-state and interstate basis.

4C. A QUANTITATIVE PERSPECTIVE.

One might ask to what extent do the type of instrument difficulties identified above influence student performance? The following section examines this question from two perspectives.

First, we will look at the issue of instrument adequacy as a whole by examining student performance in the instruments' primary domains of literacy and numeracy, and comparing the performance of Indigenous and non-Indigenous students. Gender variations in performance will be presented as a standard of comparison. The purpose of this approach is to show whether 'indigeneity' as a category discriminates more or less strongly than gender as a category. This will give us some sense of whether the tests have a particular bias or skew for or against Indigenous students' performance.

Second we will investigate response patterns within subdomains of the test, and on specific items to determine if students are experiencing difficulties in the areas identified as problematic in the associated case studies.

The Question of Relative Disadvantage.

One way to assess the degree to which Indigenous students present as a disadvantaged group is to compare them to another area of disadvantage which has been identified as a priority area. One such area is gender. Currently gender inequity is seen by education systems as a priority area, with a specific focus on the literacy and numeracy performance of boys. To assess the extent to which the 2000 Year 5 tests may be seen as a source of inequitable assessment for Indigenous students, a series of two way MANOVAs were conducted examining differences in the tests' subscales in both raw and scale form for both the literacy, and numeracy items.

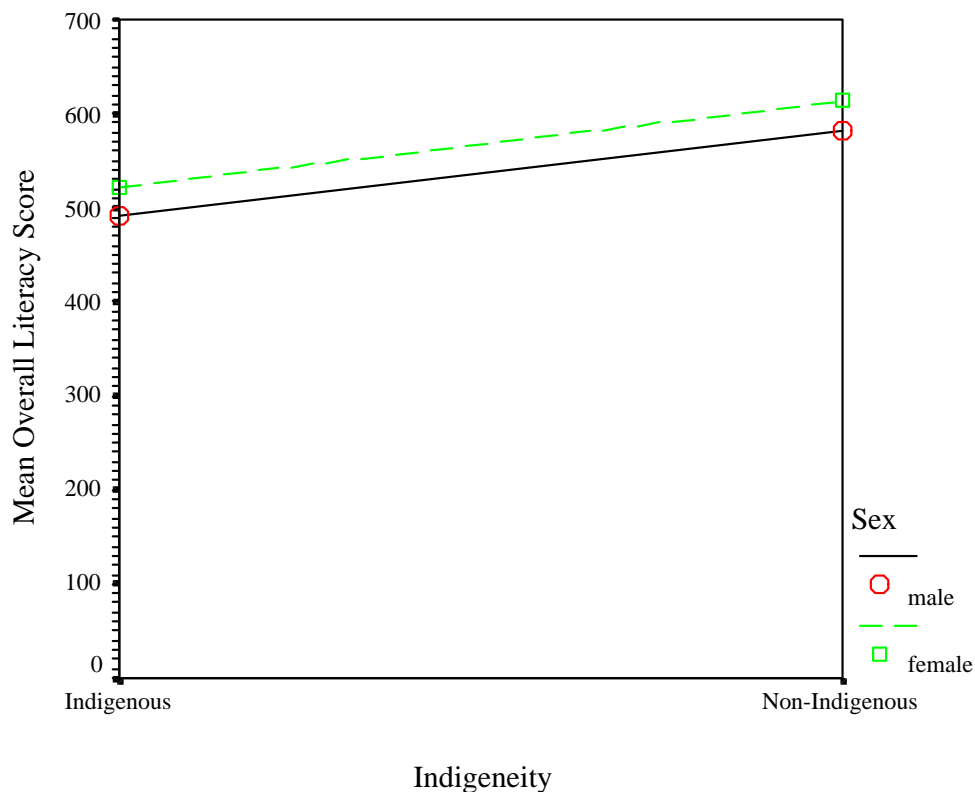
In addition, overall performance in the literacy and numeracy domains were examined by gender and indigeneity. Examination of the test effect size of the main effects of the gender and indigeneity variables provides measure of the relative impact of each variable on students' overall performance. An effect size for indigeneity at least equal to the influence of gender would suggest that indigeneity as an equity issue has an impact at least as large as gender influences.

The 2(gender) x 2(indigeneity) MANOVA analysis of overall literacy and numeracy score revealed significant main effects for both gender ($F(2, 38508) = 174.95, p < .001, \eta^2 = .009$) and indigeneity ($F(2, 38508) = 1005.54, p < .001, \eta^2 = .05$). The result indicates that indigeneity accounts for five times as much variance as difference in gender. A significant multidimensional sex by indigeneity interaction was also observed ($F(1, 38508) = 7.32, p < .01, \eta^2 = .00$). Univariate analyses revealed the interaction to be limited to numeracy ($F(1, 38508) = 5.570, p < .05, \eta^2 = .00$). Power for all tests exceeded 0.9. Figure 1 (below) shows the pattern of variation in literacy and numeracy scores by indigeneity and gender.

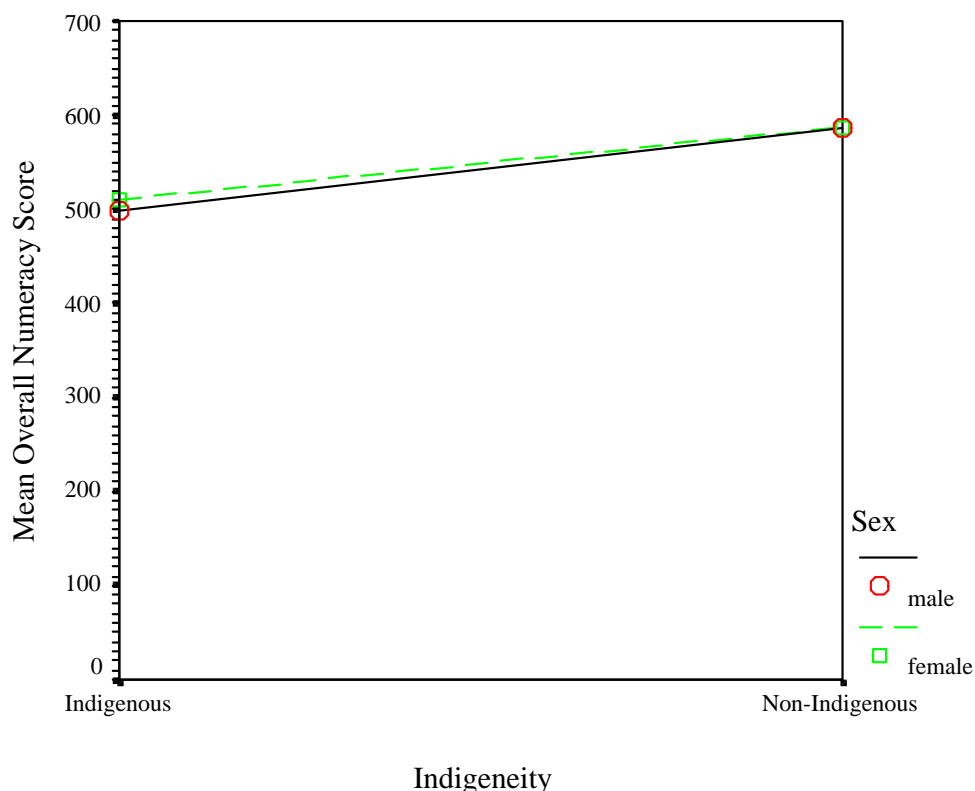
In summation the MANOVA reveals a significant and reliable, albeit small, difference in student performance by gender, and indigeneity. Moreover, the analysis reveals that test scores are much more heavily impacted upon by indigeneity than gender. Of course having said that it must be acknowledged that the overall level of disadvantage for both groups appears small. In the case of gender 99% of test performance variance is accounted for by other factors.

Figure 1: 2000 Year 5 Tests in Aspects of Literacy and Numeracy patterns of variation of response by indigeneity and gender

Literacy



Numeracy



Aspects of Literacy Instrument domain analysis.

Content

The Year 5 state-wide literacy tests measure student performance across 4 subdomains of literacy; dictation, proof-reading, reading and viewing, and writing. Individual item data is available for the first three of these sub-domains in the form of individual student responses to items, whereas only a total score is available for writing. Discriminant Function Analysis³ was used to determine to what extent the domains differentially discriminated between Indigenous and non-Indigenous respondents. It is apparent from Table 1 (below) that the Year 5 literacy tests significantly discriminate between Indigenous and non-Indigenous students on each of the domains. Whilst the magnitude of Wilk's lambda⁴ indicates that the discriminatory power of the tests is weak, this is consistent with the MANOVA result indicated above which suggested that indigeneity alone accounts for no more than 5% of the variance in test results. Moreover, if the Year 5 tests were perfectly equitable

³ Discriminant Function Analysis is a technique used to generate a weighted linear function to discriminate between two or more groups. The resultant function allows us to identify; the relative discriminating power of the set of predictors included in the analysis, the overall power of the discriminant function to discriminate between groups, and assess the efficacy of the derived discriminant function to discriminate between groups.

⁴ Wilk's Lambda is a metric of group variation. Values approaching 0 indicate that group means differ, whereas values approaching 1 represent little difference in group means relative to individual variation.

one would ideally find no significant discrimination between test respondents on the basis of indigeneity.

As the obtained Wilk's λ consistently approaches 1 on each of the Year 5 tests sub-domains it is clear that the test does not strongly discriminate between Indigenous and non-Indigenous students. However, the overall significance of summary statistics is consistent with the notion that some items on the test may be contributing to significant discrimination between groups. The size of the data set contributes to high test sensitivity which allows us to detect very small but consistent differences across the sample.

Table 1: Discriminant function analysis summary for the four literacy sub domains of the 2000 Year 5 Aspects of Literacy Test

| Subdomain | Wilk's λ | χ^2 |
|---------------------|------------------|------------------------------|
| Dictation | .985 | $\chi^2(9) = 40.56^{***}$ |
| Proof-reading | .984 | $\chi^2(7) = 623.32^{***}$ |
| Reading and viewing | .930 | $\chi^2(42) = 2710.62^{***}$ |
| Writing | .963 | $\chi^2(2) = 1406.75^{***}$ |

Table 1 suggests that the greatest discrepancy between respondents is found in the reading and viewing aspects of the Year 5 test (the smaller Wilk's λ the greater the power of the discriminant function to discriminate between groups).

Content appropriateness - Literacy item 15

The uniformly high Wilk's λ 's reported in Table 2 suggests that the individual reading and viewing items do not discriminate well between Indigenous and non-Indigenous students. However, the significant univariate F tests indicate that there are measurable differences in performance on some items.

There is some evidence that specific items do have a negative discriminatory effect on Indigenous students' performance. Moreover, these effects may reflect the cultural content issues raised by the panel.

For example, item 15 indicates as an item which significantly discriminates between Indigenous and non-Indigenous students. This finding is validated by the findings of the panel of Indigenous educators reported earlier which identified content issues with this item, and those around it. Item 15 relates to dog ownership, with the subtext being that pets should not be treated as toys. Whilst 86.7% of non-Indigenous children identified the desired response (that puppies cannot be discarded like toys), only 59.9% of Indigenous children chose that response (refer to Table 3 below). Notice that the pattern of responses by Indigenous children does not seem to be uniformly distributed across items suggesting that the Indigenous students' responses are considered (hypothetically, on the basis of culture-specific knowledge and experience) and not random.

Table 2: Analysis of reading and viewing items from the 2000 Year 5 Aspects of Literacy Test sorted by discriminating power

| Item | Wilks' Lambda | F | Item | Wilks' Lambda | F |
|------|---------------|----------|------|---------------|---------|
| 15 | 0.966 | 1304.044 | 38 | 0.988 | 449.327 |
| 11 | 0.974 | 993.298 | 22 | 0.989 | 406.242 |
| 10 | 0.977 | 878.646 | 24 | 0.990 | 388.124 |
| 40 | 0.977 | 871.639 | 42 | 0.990 | 391.477 |
| 9 | 0.980 | 753.025 | 5 | 0.991 | 319.9 |
| 29 | 0.980 | 751.1 | 18 | 0.991 | 347.568 |
| 8 | 0.982 | 666.293 | 30 | 0.991 | 323.536 |
| 23 | 0.983 | 636.269 | 2 | 0.992 | 303.54 |
| 31 | 0.984 | 589.954 | 41 | 0.992 | 288.944 |
| 32 | 0.984 | 619.941 | 35 | 0.993 | 246.134 |
| 39 | 0.984 | 618.28 | 7 | 0.994 | 211.626 |
| 33 | 0.985 | 567.999 | 27 | 0.994 | 225.88 |
| 25 | 0.986 | 541.369 | 36 | 0.994 | 212.673 |
| 26 | 0.986 | 548.051 | 6 | 0.995 | 199.828 |
| 1 | 0.987 | 505.482 | 16 | 0.996 | 159.012 |
| 3 | 0.987 | 485.135 | 21 | 0.997 | 116.335 |
| 4 | 0.988 | 445.144 | 13 | 0.998 | 76.075 |
| 12 | 0.988 | 469.097 | 34 | 0.998 | 88.543 |
| 17 | 0.988 | 448.982 | 14 | 0.999 | 40.41 |
| 28 | 0.988 | 449.396 | 19 | 0.999 | 49.76 |
| 37 | 0.988 | 452.887 | 20 | 0.999 | 30.098 |

Table 3: 2000 Year 5 Aspects of Literacy Test Item 15 response patterns broken down by indigeneity

Correct response: Puppies cannot be thrown away like toys

| | Indigeneity | | Total |
|--|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| Rubbish bins are dangerous for puppies | 13.2% | 5.8% | 6.4% |
| Naughty puppies belong in rubbish bins | 6.5% | 1.5% | 1.9% |
| Puppies cannot be thrown away like toys | 59.9% | 86.7% | 84.6% |
| Rubbish bins can be used as safe kennels for new puppies | 8.5% | 2.7% | 3.1% |
| Multiple response | .1% | .1% | .1% |
| Missing data | 11.8% | 3.2% | 3.9% |
| Total | 100.0% | 100.0% | 100.0% |

Selection of Text - Literacy Items 17-21

The panel of Indigenous educators identified a general concern with the appropriateness of the content of some items relating to such issues as uncommon subject matter. This is consistent with the identification of problematic items revealed in the quantitative analysis. Item 17 for example may be considered problematic for any Year 5 respondent in as much as their domain knowledge about the research process, motivations for citing sources of evidence, and the work practices of journalists is likely to be impoverished. This appears to be reflected in the pattern of errors presented. Note that the pattern of responses to distractors (particularly for Indigenous students) is relatively uniform suggesting that students are using a uniformed guessing strategy to respond. Non-Indigenous students seem to be using a somewhat better strategy in as much as they may be excluding one response category (the possibility that the reporter works with the researcher).

Table 4: 2000 Year 5 Aspects of Literacy Test Item 17 response patterns broken down by indigeneity

Correct response: He did the research

| % within Indigeneity | | | |
|--|-------------|----------------|--------|
| | Indigeneity | | Total |
| | Indigenous | Non-Indigenous | |
| He did the research. | 44.2% | 66.4% | 64.7% |
| He wrote the article. | 16.0% | 10.9% | 11.3% |
| The reporter works with him. | 11.3% | 4.8% | 5.3% |
| The reporter agrees with his position. | 15.9% | 13.7% | 13.9% |
| multiple entries | .1% | .1% | .1% |
| Missing data | 12.4% | 4.1% | 4.7% |
| Total | 100.0% | 100.0% | 100.0% |

Test Development - Items 19 - 21

Items 19 through 21 are also problematic from a test development perspective. More than 70% of Indigenous students and 60% of non-Indigenous students chose a distractor rather than the desired response (refer to Table 5). Moreover Indigenous and non-Indigenous students alike tend to respond uniformly across the distractors suggesting that their primary response strategy is uniformed guessing. As noted above the use of content that reflected a more common animal may well be a more appropriate context. The choice of a Quoll may be justified superficially on the grounds of cultural appropriateness or providing a meaningful local context for the item. However, this would only apply if the content matter is sufficiently common to;

- be transparent to students and thus not impose any additional cognitive load,
- reduce item load by providing cues that enable a more transparent mapping of the problem specification.

Neither appears to apply here, as neither group of students benefits from the context. In fact the Indigenous students, who one might suppose should benefit from

appropriate contextualisation, seem to perform slightly worse than non-Indigenous students in as much as they are more prone to choose the distractor items.

Table 5: 2000 Year 5 Aspects of Literacy Test Items 19 - 21 response patterns broken down by indigeneity

Item 19 response patterns by indigeneity

% within Indigeneity

| | Indigeneity | | Total |
|-----------|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| Incorrect | 76.0% | 69.5% | 70.0% |
| Correct | 24.0% | 30.5% | 30.0% |
| Total | 100.0% | 100.0% | 100.0% |

Item 20 response patterns by indigeneity

% within Indigeneity

| | Indigeneity | | Total |
|-----------|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| Incorrect | 79.7% | 75.0% | 75.3% |
| Correct | 20.3% | 25.0% | 24.7% |
| Total | 100.0% | 100.0% | 100.0% |

Item 21 response pattern by indigeneity

% within Indigeneity

| | Indigeneity | | Total |
|-----------|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| Incorrect | 72.4% | 62.0% | 62.8% |
| Correct | 27.6% | 38.0% | 37.2% |
| Total | 100.0% | 100.0% | 100.0% |

Literacy Items 28 – 32

The items relating to 'viewing' as a component of Year 5 literacy were identified as particularly problematic by the panel of Indigenous educators. Items 28 through to 32 purport to assess students' abilities and skills in the viewing aspects of literacy. Quite apart from the contextual (in)appropriateness of the illustration the confounding of the domain of viewing with reading raises questions as to the validity of the construct and how to interpret student responses. It was argued that these items might prove particularly problematic to Indigenous participants from rural settings. Discriminant analysis certainly supports the view that the response pattern of Indigenous participants differs from that of non-Indigenous participants (refer to Table 2 above).

Table 6: 2000 Year 5 Aspects of Literacy Test Item 29 response patterns broken down by indigeneity

Correct response: He is the most prominent feature of the cover

Item 29: What clue suggests the boy could be the main character? Response breakdown by Indigeneity.

% within Indigeneity

| | Indigeneity | | Total |
|--|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| He is carrying a pet duck. | 36.4% | 25.7% | 26.6% |
| He is coloured in dark shades | 6.9% | 3.2% | 3.5% |
| He is the only character on the cover. | 9.5% | 6.2% | 6.5% |
| He is the most prominent feature of the cover. | 32.7% | 60.6% | 58.5% |
| multiple entry | .2% | .0% | .1% |
| Missing | 14.3% | 4.2% | 5.0% |
| Total | 100.0% | 100.0% | 100.0% |

A closer examination of the response pattern to item 29 (see Table 6 above) within Indigenous versus non-Indigenous groups indicates that there is a significantly different response pattern for Indigenous and non-Indigenous students ($\chi^2(5)=1206.614, p<.001$) Indigenous Year 5 students were more likely to choose the first distractor than the desired response, whereas 60.6% of non-Indigenous students were likely to choose the desired response. Such a pattern suggests that this may be something other than student ‘error’ per se but instead indicative of a particular response based on specific group membership.

The first distractor appears to be a preferred item even amongst non-Indigenous respondents who responded incorrectly on this item. A straightforward interpretation of this is simply that the option describes the strongest visual element in the image from a different perspective. This option may be considered equivalent in meaning to the preferred option, and represent either a naive response (that is ignorance of the visual genre) or perhaps a linguistic effect. If students are unfamiliar with the term *prominent* but have a naive understanding of the picture genre then the first option may be considered the best response. The non-random pattern of responses overall and the preference for the first distractor across both Indigenous and non-Indigenous students responding incorrectly to the item is consistent with the notion that students are not simply guessing at a response but rather applying some systematic reasoning to select this item.

Literacy Item 35

The pattern of responses to item 35 are likewise markedly different for Indigenous and non-Indigenous respondents (see Table 7 below). As with other items the rate of non-participation in the item is indicated by the proportion of missing data reported by each group for this item. The Indigenous group is three times less likely to

complete this item than the non-Indigenous group. Further examination of the pattern of responses indicates that whereas the non-Indigenous respondents have a clear response preference for item 2 (powerlifting), this preference is much less evident for Indigenous respondents. Examination of Table 7 below shows the pattern of responses for Indigenous students to be much more uniform indicating that more Indigenous students are resorting to guessing strategies.

Further evidence of a linguistic component to this pattern is evident if Indigenous students' responses are further disaggregated by NESB or parental language use. Both of which reveal little discrimination between responses when the child comes from a NESB or if the child's parent uses a language other than English at home and the child has identified themselves as Indigenous. The observation that non-Indigenous children with NESB backgrounds appear to respond similarly to non-Indigenous children with English as a first language suggests that the results of the Indigenous students reflect a specific language difficulty associated with Torres Strait Creole.

Table 7: 2000 Year 5 Aspects of Literacy Test Item 35 response patterns broken down by indigeneity

Correct response: Powerlifting

Item 35: Which of the following sports will be part of the year 2000 paralympics but not the year 2000 Olympics. Breakdown of response pattern by indigeneity.

% within Indigeneity

| | Indigeneity | | Total |
|-------------------|-------------|----------------|--------|
| | Indigenous | Non-Indigenous | |
| horse-riding | 18.6% | 14.7% | 15.0% |
| powerlifting | 29.7% | 46.9% | 45.6% |
| cycling | 19.4% | 13.8% | 14.2% |
| sailing | 12.9% | 13.1% | 13.1% |
| multiple response | 3.1% | 5.8% | 5.6% |
| missing | 16.2% | 5.7% | 6.5% |
| Total | 100.0% | 100.0% | 100.0% |

Table 8: 2000 Year 5 Aspects of Literacy Test Item 35 Indigenous students' response patterns broken down by NESB

Indigenous students responses to item 35 broken down by NESB ^a

% within Non-English Speaking Background?

| | Non-English Speaking Background? | | Total |
|--------------|----------------------------------|--------|--------|
| | Yes | No | |
| Horse-riding | 25.3% | 22.5% | 23.2% |
| Powerlifting | 31.0% | 38.7% | 36.8% |
| Cycling | 28.0% | 22.7% | 24.0% |
| Sailing | 15.6% | 16.0% | 15.9% |
| Total | 100.0% | 100.0% | 100.0% |

^a. Indigeneity = Indigenous

Table 9:2000 Year 5 Aspects of Literacy Test Item 35 non-Indigenous students' response patterns broken down by NESB

Non-indigenous responses to item 35 broken down by NESB. ^a

% within Non-English Speaking Background?

| | Non-English Speaking Background? | | Total |
|--------------|----------------------------------|--------|--------|
| | Yes | No | |
| Horse riding | 19.7% | 16.3% | 16.7% |
| Powerlifting | 47.8% | 53.5% | 53.0% |
| Cycling | 18.7% | 15.2% | 15.6% |
| Sailing | 13.9% | 14.9% | 14.8% |
| Total | 100.0% | 100.0% | 100.0% |

^a. Indigeneity = Non-Indigenous

Moreover it appears that there are district differences in response patterns. In particular if the overall response is broken down by district it becomes apparent that responses to this item actually reverses in some of the more remote districts. That is in those districts indicated by bold type in Table 10 (below), more students get the item wrong than right. These districts are typically rural/remote.

This provides us with some evidence that specific items may again discriminate not only on the basis of indigeneity per se, but as well have content that favours those students in closer proximity to urban centres (and, therefore, mainstream cultural knowledges and texts of particular types). The relationship between ethnicity, class, and location/space is, of course, a major topic of empirical and theoretical work in the field of social geography (Harvey, D. 2001).

Aspects of Numeracy domain analysis

Table 11 (below) reveals significant discrimination between Indigenous and non-Indigenous students on the individual items which constitute the numeracy domain. As noted earlier the magnitude of Wilk's lambda indicates that the discriminatory power of the tests is weak. However, it is also readily apparent that the greatest discrimination between Indigenous and non-Indigenous students is found in the three items identified as having complex literacy load earlier in this text (items 12 - 14). Item 13 is in fact the greatest single discriminator between the two groups.

Supporting information collected from the Indigenous educators, and further developing the qualitative analysis of the numeracy item 14, Table 12 (below) reports the response patterns to this item broken down by indigeneity. As the 2000 Year 5 data contains simple binary scoring for this item it is not possible to identify the characteristic answers given by Indigenous children to this set of items. However, an examination of students response patterns to this item reveals that Indigenous students produce significantly more incorrect answers than non-Indigenous students ($z = 10.08$, $p < .001$). Moreover, Indigenous students are much more likely to not-respond to the item (6.7% cf 2.5%) this represents a 168% increase in non-response for the

Indigenous students. When non-responses (missing values) are excluded from the data the difference becomes even larger.

Table 10: 2000 Year 5 Aspects of Literacy Test Item 35 raw performance of all students by district

| Raw performance by district | | | |
|------------------------------------|-----------|--------------------|------------------|
| % within DISTRICT | | | |
| | | Item 35 | |
| | | Incorrect response | Correct response |
| DISTRICT | | | |
| | BA | 49.6% | 50.4% |
| | BU | 49.8% | 50.2% |
| | CC | 59.9% | 40.1% |
| | CH | 47.6% | 52.4% |
| | CP | 53.0% | 47.0% |
| | CR | 45.8% | 54.2% |
| | DA | 51.5% | 48.5% |
| | EM | 50.2% | 49.8% |
| | FC | 53.3% | 46.7% |
| | GE | 48.7% | 51.3% |
| | GL | 50.9% | 49.1% |
| | GN | 53.4% | 46.6% |
| | GS | 52.5% | 47.5% |
| | IB | 54.9% | 45.1% |
| | IP | 59.2% | 40.8% |
| | LB | 59.6% | 40.4% |
| | LO | 50.6% | 49.4% |
| | MD | 52.9% | 47.1% |
| | MG | 46.4% | 53.6% |
| | MH | 59.2% | 40.8% |
| | MI | 63.3% | 36.7% |
| | MN | 55.4% | 44.6% |
| | MO | 52.8% | 47.2% |
| | NA | 56.2% | 43.8% |
| | RH | 54.3% | 45.7% |
| | RO | 57.0% | 43.0% |
| | SB | 58.9% | 41.1% |
| | ST | 43.3% | 56.7% |
| | TB | 55.4% | 44.6% |
| | TJ | 55.9% | 44.1% |
| | TN | 56.9% | 43.1% |
| | TO | 47.7% | 52.3% |
| | TS | 68.7% | 31.3% |
| | WA | 52.9% | 47.1% |
| | WM | 54.6% | 45.4% |

Table 11: 2000 Year 5 Aspects of Numeracy Test - Discriminant function analysis summary

| Tests of Equality of Group Means | | | | | |
|----------------------------------|---------------|--------|-----|------|------|
| | Wilks' Lambda | F | df1 | df2 | Sig. |
| Y5NUMR1 | .998 | 4.312 | 1 | 2762 | .038 |
| Y5NUMR2 | .998 | 4.615 | 1 | 2762 | .032 |
| Y5NUMR3 | .998 | 6.734 | 1 | 2762 | .010 |
| Y5NUMR4 | .998 | 5.281 | 1 | 2762 | .022 |
| Y5NUMR5 | 1.000 | .426 | 1 | 2762 | .514 |
| Y5NUMR6 | .999 | 3.522 | 1 | 2762 | .067 |
| Y5NUMR7 | .997 | 9.617 | 1 | 2762 | .002 |
| Y5NUMR8 | .997 | 8.490 | 1 | 2762 | .004 |
| Y5NUMR9 | .997 | 7.703 | 1 | 2762 | .006 |
| Y5NUMR10 | .999 | 4.126 | 1 | 2762 | .042 |
| Y5NUMR11 | 1.000 | 1.119 | 1 | 2762 | .290 |
| Y5NUMR12 | .993 | 18.365 | 1 | 2762 | .000 |
| Y5NUMR13 | .993 | 20.012 | 1 | 2762 | .000 |
| Y5NUMR14 | .994 | 17.459 | 1 | 2762 | .000 |
| Y5NUMR15 | .994 | 17.195 | 1 | 2762 | .000 |
| Y5NUMR16 | .998 | 5.297 | 1 | 2762 | .027 |
| Y5NUMR17 | .994 | 18.063 | 1 | 2762 | .000 |
| Y5NUMR18 | .996 | 11.321 | 1 | 2762 | .007 |
| Y5NUMR19 | .997 | 8.722 | 1 | 2762 | .003 |
| Y5NUMR20 | .997 | 8.992 | 1 | 2762 | .003 |
| Y5NUMR21 | .996 | 10.794 | 1 | 2762 | .007 |
| Y5NUMR22 | .997 | 8.819 | 1 | 2762 | .003 |
| Y5NUMR23 | 1.000 | .239 | 1 | 2762 | .625 |
| Y5NUMR24 | .997 | 7.007 | 1 | 2762 | .008 |
| Y5NUMR25 | .997 | 7.333 | 1 | 2762 | .007 |
| Y5NUMR26 | .998 | 5.314 | 1 | 2762 | .027 |
| Y5NUMR27 | .999 | 2.198 | 1 | 2762 | .138 |
| Y5NUMR28 | .997 | 7.126 | 1 | 2762 | .008 |
| Y5NUMR29 | 1.000 | .000 | 1 | 2762 | .984 |
| Y5NUMR30 | .999 | 3.288 | 1 | 2762 | .070 |
| Y5NUMR31 | .999 | 3.581 | 1 | 2762 | .059 |
| Y5NUMR32 | .997 | 7.773 | 1 | 2762 | .005 |
| Y5NUMR33 | .998 | 6.372 | 1 | 2762 | .012 |
| Y5NUMR34 | 1.000 | .792 | 1 | 2762 | .374 |
| Y5NUMR35 | 1.000 | .482 | 1 | 2762 | .487 |
| Y5NUMR36 | .998 | 5.597 | 1 | 2762 | .018 |
| Y5NUMR37 | .998 | 5.093 | 1 | 2762 | .024 |
| Y5NUMR38 | 1.000 | .059 | 1 | 2762 | .808 |
| Y5NUMR39 | .999 | 3.998 | 1 | 2762 | .046 |
| Y5NUMR40 | 1.000 | .147 | 1 | 2762 | .707 |
| Y5NUMR41 | 1.000 | .213 | 1 | 2762 | .644 |
| Y5NUMR42 | .997 | 7.557 | 1 | 2762 | .006 |
| Y5NUMR43 | 1.000 | .331 | 1 | 2762 | .565 |
| Y5NUMR44 | 1.000 | .226 | 1 | 2762 | .634 |
| Y5NUMR45 | 1.000 | 1.108 | 1 | 2762 | .293 |

Table 12: 2000 Year 5 Aspects of Numeracy Test Item 14 response patterns broken down by indigeneity

| Indigeneity | | Frequency | Percent | Valid Percent |
|-----------------------|-----------------------|------------------|----------------|----------------------|
| Indigenous | Incorrect | 2212 | 74.7 | 80.0 |
| | Correct | 552 | 18.6 | 20.0 |
| | Total valid responses | 2764 | 93.3 | 100.0 |
| | Missing | 198 | 6.7 | |
| | | 2962 | 100.0 | |
| Non-Indigenous | Incorrect | 23360 | 65.7 | 67.4 |
| | Correct | 11317 | 31.8 | 32.6 |
| | Total valid responses | 34677 | 97.5 | 100.0 |
| | Missing | 892 | 2.5 | |
| | | 35569 | 100.0 | |
| Unknown | Incorrect | 300 | 42.5 | 76.9 |
| | Correct | 90 | 12.7 | 23.1 |
| | Total valid responses | 390 | 55.2 | 100.0 |
| | Missing | 316 | 44.8 | |
| | | 706 | 100.0 | |

This would appear to verify the panel’s claim that the numeracy items which require higher and more complex levels of reading comprehension discriminate against Indigenous students, generally more likely to be EFL/ESL/ESD than the non-Indigenous population. Further it suggests that issues of construct validity in literacy-contingent numeracy items needs further systematic investigation.

4D. SUMMARY

To summarise then, the foregoing analysis was undertaken in three stages, moving from (1) expert Indigenous educational judgements about tests and test items, to (2) closer linguistic and content analysis of items of the types adjudged to be potentially discriminatory; to (3) statistical analysis of overall test results and response patterns to these same items to check for evidence of cultural and/or linguistic discrimination and bias.

The findings here are illustrative and exemplary, limited by the scope and duration of the present research project. A comprehensive, item-by-item analysis of the Years 3, 5 and 7 Aspects of Literacy and Numeracy tests would be needed in order to make overall claims about which items, sections and modes have discriminatory effects against Indigenous students. This would give us a total picture. However, we can offer the following findings from this analysis:

- That some literacy and numeracy items on the 2000 Queensland Year 5 tests have discriminatory effects on the basis of language demands and culture-specific content knowledge;

- That there are some clear unresolved issues in construct validity on several of the test items;
- That the framing and contexts of test administration requires continued careful scrutiny.

These findings are not surprising and could have been predicted solely on the basis of expert Indigenous opinion and field-based interviews in the few schools sampled. *What is remarkable is that there have been no prior such published studies or systematic investigation of assessment-bias in Queensland or Australia in the last decade, despite concerns about culture-bias in standardised testing raised internationally for the past four decades and raised by Queensland psychologists and educational researchers as early as 1970.*

The power of this research study is not just in these limited findings. We believe that the panel/ linguistic item analysis/statistical analysis procedures undertaken here could become routine, extremely cost-efficient components of test development and analysis by the QSA, Education Queensland and their research and development contractors. The results would not be perfect – but we believe that the results could be *more* culture-fair standardised achievement tests, that are able to more accurately and more fairly assess and report on Indigenous students' performance in literacy and numeracy. It is important to make clear that this isn't a matter of 'dumbing down' standards – but of working in more expert, technically sophisticated and scientific ways.

We conclude with an overall analysis of the issues around culture-fair assessment and reporting, key findings and recommendations.

5: POLICY ANALYSIS ISSUES FOR APPLICATIONS OF STATE-WIDE ACCOUNTABILITY MEASURES

This report is not anti-testing. It supports the development of culture-fair, technically rigorous, and, within their own technical limitations, educationally useful assessment, evaluation and testing. We offer a series of very concrete, cost-effective ways of both improving the current Year 2 Diagnostic Net and its planned replacement, and the current and any planned state-wide tests. We do not here wish to blame or accuse either the tests or their development processes for omissions or problems – any of which, we have noted here, have fallen purely within the conventions of an overall neglect by Australian educators, systems and agencies, of issues related to culture-fairness in the increasingly important instrumentation for assessing student and system outcomes.

Our approach to the task here was very straightforward and rudimentary – however technically complex its language and argumentation have been:

- We reviewed the international and national literature on testing bias and culture-fair testing and evaluation, using this literature to set up a framework of questions for examining and classifying our examinations of current instrumentation;
- We reviewed the histories, contexts and development of the major instruments in Queensland: The Year 2 Diagnostic Net, the Years 3, 5 and 7 Aspects of Literacy and Numeracy tests;
- We made field visits to selected schools to gather information on how this instrumentation was used and results reported;
- We convened an expert panel to evaluate illustrative test data, items and findings for cultural and linguistic bias;
- We undertook more detailed linguistic, content and statistical analysis of selected items.

In this final section, we offer a series of findings and recommendations in the context of a general thematic discussion of the strengths and weaknesses of both the Year 2 Diagnostic Net and the State-wide Aspects of Literacy and Numeracy testing processes.

5A. POSITIVE BY-PRODUCTS OF THE YEAR 2 DIAGNOSTIC NET PROCESS

There are significant positive features of the Year 2 Diagnostic Net that have productive consequences for Indigenous students and their teachers. Simply, the Net

has provided a face-to-face, point-in-time assessment of individual students as they begin to engage with early literacy. It has provided teachers the opportunity to engage directly with aspects of Indigenous students' language and literacy development. In the case studies reported here and in the existing literature, we have found that there is evidence that the Year 2 Diagnostic Net has:

- provided a common language for understanding of literacy and numeracy acquisition;
- developed teacher expertise in linking and aligning pedagogy and assessment strategies;
- assisted schools in developing effective systems of monitoring student progress over time;
- provided a common language for the matching of performance against standards and targets;
- provided and supported the development of useful networks of early childhood teachers across schools, and throughout the State; and
- standardised the systematic moderation procedures for the Net around a view of appropriate literacy and numeracy targets.

The Year 2 Diagnostic Net (Net) has provided a powerful professional development forum. For those who argue that standardised achievement tests are by definition discriminatory in their format and administration – the face-to-face diagnostic assessment of the Net, its moderation procedures and reporting system has provided a valuable alternative. In several of the case study schools we visited, the Net provided a forum for talk and discussion about Indigenous students' progress.

It is also noteworthy that in one non-government school that we visited, that did not have the Net in place, there were real problems with early assessment and tracking of Indigenous children's performance, with the learning support/special education teacher charged with using an unnormed, dated Neale Reading Test for early diagnostic purposes. This situation is - our field-visits and expert panel informed us - all too often the case across the state.

There is however a critical problem with the Net. As we noted in the introduction, one of the principal ways in which assessment and evaluation can discriminate against culturally different students is the extent to which it tends to concretise or formalise monocultural, singular descriptions of sequences of child development (McNaughton, 2002). This aspect of the Year 2 Diagnostic Net and the *First Steps* Continua has been raised repeatedly in the literature but never systematically investigated.

At the time of this report writing, the current QSCC has announced its intention to 'roll' the Net procedures into the new English syllabus currently under trial. Nonetheless, we recommend yet again that the Net - even in its newly conceived form - be rigorously reviewed and modified to accommodate alternative, Aboriginal and Torres Strait Islander patterns and pathways of language and literacy development. By drawing upon current developments in other States, specifically revisions to *First Steps* underway in Western Australia, this could be done without a major research project. The principal cost might involve the development of alternative reading text materials with different cultural and linguistic content. However systematic

investigation of any tool being used as a systemic assessment of literacy development must be rigorous and continuous.

5B: LIMITATIONS OF THE STATE-WIDE TESTS

As noted, the expert panel was sceptical of the format, content and linguistic presuppositions of the Years 3, 5 and 7 tests. However, they did not object to testing per se. Quite the contrary, the Indigenous principals we spoke to argued that it gave them some clear baseline data around which to set targets for planning and intervention. This was particularly important in those community schools, where the transitory nature of staffing made whole school planning, targets and diagnostic work all the more important.

By contrast, the teachers we spoke to in the field generally perceived the Years 3, 5 and 7 tests as largely external, mistimed, problematic to administer to Indigenous students and more difficult to use as formative or diagnostic purposes for classroom intervention.

Some of the problems with the current testing processes include:

- the point in time nature of the tests;
- the November reporting date which precludes any useful linkage into classroom planning and teaching for that year (a particular issue for Year 7 students given their transitions into many secondary contexts);
- the lack of processes and funding in place to encourage teacher networking and moderation procedures or even appropriate interpretation of test data by administrators and teachers;
- the formal and technical nature of the reports to Aboriginal and Torres Strait Islander parents which are not typically mediated by schools or classroom teachers, but the results of which teachers are still required to be accountable for. The issue of how to accurately and constructively report student results and progress to Aboriginal and Torres Strait Islander parents, families, and Elders is a persistent problem in the field and will require serious research and development if the issue is to be solved;
- the tests' narrow coverage of domains of literacy and numerary, while still attempting to reflect the complex modal varieties of the syllabus;
- the fact that the tests attempt to provide accountability information at several levels simultaneously - at systemic, school-based and individual student levels.

In summary, the external nature of the tests and the formal reporting procedures work toward limiting teacher expertise and involvement in the testing procedures. These generic limitations are further exacerbated when the tests are used to assess Indigenous students, particularly with EFL/ESL/ESD learners in a range of urban and rural contexts as well as those located in more isolated and/or community contexts. The findings of our cases study of the 2000 Year 5 tests were corroborated by our interviews in the field. These identified:

- **Language issues:** The tests purport to assess domains related to literacy and numeracy competencies, and yet are in fact tests of competence in Standard Australian English, with considerable confusions related to exemptions, special considerations and support provisions for EFL/ESL/ESD learners in different Indigenous community contexts;
- **Cultural content issues:** Sampling of presumed common life experiences within test items can significantly disadvantage Indigenous students in remote and/or community contexts given their limited direct contact with such experiences;
- **Framing issues:** Test administration procedures and practice sessions can generate considerable anxieties for both teachers and students, which in combination can affect both overall attendance by Indigenous students and quality test preparation organised by their teachers;

5C. INSTITUTIONALISATION OF THE NET AND TESTS AS SYSTEMIC ACCOUNTABILITY MEASURES

The State has accepted that Commonwealth funding is contingent on agreement to full cohort testing at each of Years 3, 5 and 7 in aspects of literacy and numeracy, and on public reporting against national literacy and numeracy benchmarks. The inter-governmental agreement and significant funding places these testing initiatives at the core of the State's approach to accountability for standards of literacy and numeracy. But such accountability practices have been applied in the State and non-State sectors in quite different ways.

For Education Queensland, the combination of the Net and tests are cornerstones from which many other initiatives flow - in the first instance intervention funding and specialist staffing based on incidence figures. Increasingly, it would appear that performance patterns over time on these accountability measures are being used as key criteria for judging effective school administration procedures, teaching quality at district and school levels and ultimately the relative overall success of state schools in raising literacy and numeracy standards at a systemic level. In the absence of agreement over other systemic indicators of outcome performance, the Net and the Tests in combination have become reified, generic indicators of quality outcomes way beyond both their original purposes and their present validity.

Whatever their histories and limitations, the Net and test data have become the most important key performance indicators for schools in their annual reporting to systems and stakeholders. If this is the case, for them to proceed without more systematic attention to issues of culture-fairness is a serious problem for Aboriginal and Torres Strait Islander communities.

The performance databases compiled by Education Queensland for each State school record overall student performance on one frame and that of Indigenous students on

another. At a broader structural level, therefore, the marginalisation process exemplified in test items is further institutionalised through public reporting processes that inaccurately group all Indigenous students as if they are a homogenous category of learners. Such a category system can reinforce a notion that lower expectations of literacy and numeracy performance is somehow to be expected among most or even all Indigenous students, based on a continued pattern of poor performance recorded over several years on the same flawed instruments.

For schools with significant numbers of Indigenous students, the institutionalisation of these accountability measures presents many and complex dilemmas in establishing priorities for school-wide policies and focal points for individual classroom teaching practices. On the one hand, the essential bias inherent in the tests and unfairness of the testing process, of specific items, and of the reporting of results for many Indigenous learners are obvious to most (Indigenous and non-Indigenous) staff in such schools as those we spoke to. On the other hand, most staff and most parents associated with these schools expect and want an education for Indigenous students that is comparable to that offered in mainstream contexts, and this includes outcome performance being measured in common ways for most students, with limited application of special conditions or exclusions.

In those schools where there are fewer Indigenous students in the enrolment mix and where they are viewed as a small minority with limited Indigenous community identification, the cultural component tends to be conveniently overlooked in the testing process while the reporting of test outcomes for Indigenous students are re-positioned as deficits within them as individuals - but still conveniently grouped by cultural identification.

There have been several commissioned reports cited here that have mentioned the issue as needing systematic study. There has also been considerable advocacy for reform of the tests, administration procedures and the application of special conditions arising from perceptions of the needs of various groups of learners - including from community members, school administrators and teachers alike in Indigenous community school settings and particularly in relation to EFL/ ESL/ESD issues. For many policy makers, it would appear that the issue is largely one of identifying those students to whom extra support or exclusion conditions should apply. Relative performance on the National ESL Bandscales, or on its EQ-developed Indigenous variation, is now taken as the critical determinant of whether or not an individual student can be considered as eligible for special consideration or exclusion from the testing process.

However, there appears to have been very limited 'root and branch' test reform undertaken by curriculum and statutory bodies in Queensland and in other States and Territories – that is, of the test item development process itself to better reflect student diversity, in particular linguistic and/or cultural difference. In essence, students in such groups are viewed as the 'exceptions that prove the rule' in a common curriculum and accountability regime. Only limited tinkering with the Net and the Tests has been undertaken over recent years, with limited budget allocations cited as the key reason. In addition, the development and piloting of the new English and Mathematics syllabuses for Years 1-10 has been underway during this time and there

have been expectations that greater alignment between curriculum and assessment systems will provide more school-based support.

In the meantime the institutionalising of the Net and tests as systemic accountability measures of increased stakes have continued apace within Education Queensland. For the non-State sector, issues related to differential take-up of the Net and tests come into play across both diocesan, systemic schools (which tend to use these measures as a matter of policy) as well as in order-owned schools in the Catholic system (which tend to reflect the more limited involvement pattern of Independent schools).

The diversity, differential governance structures and funding arrangements that characterise the non-state sector make it very difficult to generalise further about the impact of the Net and tests on either policy or practice, and for Indigenous and non-Indigenous students alike. At the very least, the extent of discretionary use of the Net and Tests in the non-State sector, particularly in Independent school contexts, makes it difficult to identify the literacy and numeracy performance of the whole cohort of students, as well as examine and analyse the comparative impact of differential support provisions for Indigenous students in a wide variety of school and community contexts.

6: CONCLUSION: A CLEAR WAY FORWARD

Our limited school-based case studies did not find generalisable models of school and classroom-based assessment and reporting. There were some excellent examples of practice in the field – but they were idiosyncratic, highly localised and would be difficult to generalise on a larger scale. The recent efforts at whole school planning as a result of *Literate Futures* (Education Queensland, 2000c), the establishment of Literacy Development Centres and IETA are beginning to draw together expertise and generalise from some exemplary practices across the state. But at present the situation is largely unchanged from what we observed in 2000 when we completed the *Literate Futures* review. The assessment practices, the overall consciousness of the specific language and literacy needs of Indigenous students, and the specific issues around culture-fair assessment are largely unaddressed throughout the system.

Some schools in the appended cases (see Appendix 1) did indeed manage to put together an effective mix of diagnostic assessment, record keeping and, occasionally, effective ways of reporting student progress to Aboriginal and Torres Strait Islander parents. If there are ‘best practices’ to be found in the field, they consisted of conscientious principals and teachers – usually a literacy coordinator working with a coordinator of Aboriginal and Torres Strait Islander students in the school – using the Net data, weighing it against testing data for trends, tracking individual students through mainstream and relevant adjunct programs (e.g., Reading Recovery, ESL intervention and learning support), and using portfolio and face-to-face reporting to parents. But such cases were rare – utterly dependent upon the highly refined training and expertise of individual teachers – and often undertaken with little if any systemic or inservice support.

In sum – in the search for best practices in the field – we found some exemplary schools and programs. But even there, they had managed to put together programs and assessment procedures that appeared to be the products of opportunity, of commitment by coordinators and individual teachers scrabbling for time, of ‘whatever was on the shelf’ in the way of standardised instruments, and out of individual experience, rather than any systematic or researched guidance from the centre or elsewhere.

What this means is that the Net and tests take on all the more importance for Aboriginal and Torres Strait Islander children, their teachers and their schools. They stand as the few systematic, relatively predictable, and perceived as valid and reliable instruments – albeit ones that are being used far beyond the scope of their original intentions or designs. *It is imperative that these be made as culture-fair as possible.*

In the present QIECB study, we have reviewed the literature. We have observed illustrative cases of the assessment of Indigenous children in the contexts of several Queensland classrooms. We have conferred with Indigenous educators and systems’ experts over the assessment instruments. We have linguistically analysed selected test

items, and we have reanalysed selective test and item data. We found that the testing and assessment currently underway at systemic and classroom level, while not deliberately discriminatory, while technically in compliance with statutory requirements, while in line with other states' practices, and while technically conventional in procedure and format, has not been designed or implemented with questions about potential cultural bias in mind.

It is worth noting again, that none of the Indigenous educators and administrators we interviewed as part of this study opposed standardised assessment. All parties agreed of the need for standards, of the need for high expectations for Indigenous learners, and that knowing where kids stand in relation to their peers across the state and across schools and systems was important. But they wanted these systems to be 'fair' in content, in administration and in reporting. And the teachers that we spoke with, Indigenous and non-Indigenous consistently reported their frustration with the quality of the data, the quality of the advice and the quality of the testing instruments that they were required to administer.

Our position is that current testing and assessment can be improved to better measure and represent what Aboriginal and Torres Strait Islander students can do in literacy, numeracy and other areas. This can be done, we believe, without psychometrically or technically compromising 'standards', validity, reliability or accuracy, and without 'dumbing down' the tests.

Quite the contrary, we argue that aspects of construct validity can be improved through adjusting test items and trialing alternative test formats and administration procedures. We do not believe that it is a matter – as was put to us by some of the educators we interviewed as part of this study – of either maintaining the current approaches and instruments or risk compromising standards, targets and expectations for Indigenous students. Better testing and assessment can be done, we believe, with some careful but not excessive investment in technical research and development.

6A. FINDINGS

- F1.** That there is some evidence that some specific test items discriminate against Aboriginal and Torres Strait Islander students on the basis of culture specific background knowledge and linguistic background.
- F2.** That there is some evidence that the test administration and reporting formats discriminate against Aboriginal and Torres Strait Islander students on the basis of culture-specific background knowledge and linguistic background.
- F3.** That efforts to extend tests to cover more complex syllabus outcomes, constructivist problem solving in numeracy, and other areas had generated a technical complexity, as well as complex issues of cognitive load and construct validity for the structure of the tests.

- F4** That there have been repeated documented calls for a critical re-evaluation of the cultural-specificity of the First Steps Developmental Continua as used on the Year 2 Diagnostic Net, and its affiliated support materials.
- F5.** Although there was some evidence in the small sample of good school-level practice, the research found that the actual school and classroom based practices for assessing students and reporting on their results tend to be, at best, highly variable and school-specific, and at worst, *ad hoc*.
- F6.** There was highly uneven evidence that system authorities had recognised that there might be problems with assessment bias and that system level attempts to adjust, accommodate or make more appropriate existing instrumentation and materials were sporadic, under-funded and had not had a major impact on practices in the field.

6B. RECOMMENDATIONS

- R1.** That the QSA, EQ, QIECB and relevant systemic authorities should establish a working group on assessment of Aboriginal and Torres Strait Islander students, with requisite, appropriately funded technical expertise in measurement *and* cultural issues. This group's charge needs to go beyond consultation and 'sign-off' on assessment instrumentation to involve the development and oversight of a proactive research and development agenda to put Queensland in the vanguard of culture-fair assessment.
- R2.** That specific achievement data should only be used for the purposes it was designed to be used for. That information about the purposes, limits, uses and abuses of specific Net and test data be circulated widely across all systems, and that critical caveats on the misuse and misinterpretation of achievement data for purposes it was not designed for be put in place. For example, non-generalisable data should not be generalised. Individualised data should not be disaggregated. Formative and diagnostic data should not be used for summative purposes. There is extensive evidence across the system of sanctioned abuse and misrepresentation of Net and test data.
- R3.** That teachers and schools working with Indigenous students should receive both system-wide advice and training in approaches to culture-fair assessment, on techniques for interpreting test data and its limits, and on techniques and approaches for reporting on student performance to parents, Elders and communities in meaningful and educationally useful ways.
- R4.** That QSA, EQ and other systems develop and trial culturally relevant materials for Aboriginal and Torres Strait Islander students, for use in face-to-face diagnostic assessment (e.g., the Net and Reading Recovery Observation Survey)
- R5.** That QSA, EQ and other systems undertake a study of whether the Developmental Continua and the outcomes statements that relate to levels 1, 2 and 3 of the Years 1-10 English Syllabus presently being trailed, are able

to accurately accommodate multiple and diverse sequences of the development of literate practices.

- R6.** That there is a need, as an experimental research and development task, to prototype and trial several alternative test items on Year 3, 5 and 7 tests that accommodate and use Indigenous knowledge.
- R7.** That the possibility of developing and deploying alternative and adjusted test administration environments and procedures for children in remote or EFL communities should be examined.
- R8.** That there is a need to re-examine the language demands and construct validity issues in the Years 3, 5 and 7 numeracy tests.
- R9.** That there should be an effort to develop simpler, less linguistically contingent and technically complex formats and item structures of assessing basic numeracy and literacy in future tests. This may require a reconsideration of efforts to extend syllabus coverage through the tests.
- R10.** That the QSA, EQ and other bodies should examine alternative categories for internal reporting and analysis of state testing results, disaggregating Aboriginal and Torres Strait Islander achievement by location, by main language experience (EFL/ESL/ESD) in order to better identify ostensive gains and losses among some sub-groups of students.

REFERENCES

- Australia. Quality of Education Review Committee. Committee chaired by Peter Karmel. (1985). *Quality of Education Review Committee*. Canberra: Australian Government Printing Service.
- Australian Council for Educational Research. (1993). *The impact of First Steps on the reading and writing ability of Western Australian Year 5 students*. Hawthorne: Australian Council for Educational Research.
- Berlack, H. (2001). Race and the Achievement Gap. *Rethinking Schools On Line*, 15 (4) [On-Line] Retrieved from World Wide Web in October 2001 from www.rethinkingschools.org.
- Blum, J.M. (1978). *Pseudoscience and Mental Ability*. New York: Monthly Review Press.
- Boulton-Lewis, G.M., Neill, H. & Halford, G.S (1986). Information processing and scholastic achievement in Aboriginal Australian children in South-East Queensland. *Aboriginal Child at School*, 14 (5), 42-55.
- Braken, B.A. & McCallum, R.S. (1998). *Universal Nonverbal Intelligence Test Examiners' Manual*. US: Riverside Publishing
- Bruner, J.S. & Haste, H. (1987). *Making Sense: The child's construction of the world*. London: Methuen.
- Callahan, R. (1962). *Education and the Cult of Efficiency*. Chicago: University of Chicago Press.
- Cambourne, B. (1988). *The whole story: Natural learning and the acquisition of learning*. Auckland: Ashton-Scholastic.
- Christie, P. (1998) Point and counterpoint: Benchmarking. *Curriculum Perspectives*, 8 (3), 43-49.
- Cole, M. (1999). *Cultural Psychology*. Cambridge: Cambridge Press.
- Commonwealth Department of Education, Training and Youth Affairs. (1997). *National Literacy and Numeracy Benchmarks*. Canberra: Department of Education, Training and Youth Affairs.
- Commonwealth Government of Australia. (2000). *The National Indigenous English Literacy and Numeracy Strategy 2000-2004*. Canberra: Commonwealth Government of Australia.
- Department of Education, Queensland. (1995). *Literacy Validation*. Brisbane: The State of Queensland.

- Department of Employment, Education and Training. (1991). *The Australian Language and Literacy Policy*. Canberra: Australian Government Printing Service.
- Department of Employment, Education and Training. (2001). *Education: National literacy and numeracy benchmark*. [On Line]. Retrieved from the World Wide Web in September 2001 at <http://www.detya.gov.au/schools/literacy7nemeracy/benchmarks.htm>.
- Deschamp, P. (1994). *The implementation of the literacy component of the First Steps project in ELAN schools*. [On Line]. Retrieved from the World Wide Web in September 2001 from http://www.ecurl.com.au/src/Firststep/Res_docs/res3.html.
- Deschamp, P. (1995). *The development and implementation of the First Steps project in Western Australia*. Perth: Department of Education.
- DEST. (2002). *The National Report on Schooling*. Melbourne: Commonwealth Government
- ECU resources for Learning. *First Steps General Information*. [On Line]. Retrieved from the World Wide Web in September 2001 from <http://www.ecurl.com.au/src/Firststep/main-firststep.html>.
- Education Queensland. (2000a). *Partners for Success. Strategy for the Continuous Improvement of Education and Employment outcomes for Aboriginal and Torres Strait Islander Peoples in Education Queensland*. Brisbane: Education Queensland.
- Education Queensland. (2000b). *Implementing Partners for Success*. Brisbane: Education Queensland.
- Education Queensland (2000c). *Literate futures: Report on the literacy review for Queensland State Schools*. Brisbane: Education Queensland.
- Fischer, C.S., Hout, M., Jankowski, M., Lucas, S.R., Swindler, A. & Voss, K. (1996). *Inequality by Design*. Princetone: Princetone University Press.
- Freebody, P. (1993). Social Class and Reading. In A. Luke & P. Gilbert (Eds). *Literacy in Contexts: Australian Perspectives and Issues* (pp. 68-84). Melbourne: Allen and Unwin.
- Footnotes. (2000). *Footnotes: A publication for First Steps tutors*. [On Line]. Retrieved from the World Wide Web in September 2001 from <http://www.ecurl.com.au/src/firststep-frameset.html>.
- Goodnow, J.J. (1976). Some sources of cultural differences in performance. In G.E. Kearney & D.W. McElwain. (Eds). *Aboriginal Cognition: Retrospect and prospect*. New Jersey: Humanities Press.

- Gould, S.J. (1981). *The Mismeasure of Man*. New York: Norton.
- Gregory, E. (1997). *One Child, Many Worlds*. New York: Teachers College Press.
- Griehaber, S. (1997). Back to the basics: The Queensland Year 2 Diagnostic Net. *Curriculum Perspectives*, 17 (3) 28-38.
- Halliday, M.A.K. (1978). *Language and Social Semiotics: The social interpretation of language and meaning*. London: E Arnold.
- Hartigan, J.A., & Wigdor, A. (Eds). (1989). *Fairness in Employment Testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington DC: National Academy Press.
- Harvey, D. (2001). *Spaces of Hope*. Edinburgh: Edinburgh University Press.
- Herrnstein, R. & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life*. New York: Free Press.
- House of Representatives Standing Committee on Employment, Education and Training. (1993). *The Literacy Challenge*. Canberra: Australian Government Printing Service.
- Kearins, J. (1983). A quotient of awareness. *Education News*, 18 (4), 18-22.
- Ladson-Billings, G. (1994). *The Dreamkeepers: Successful teachers of African-American children*. San Francisco, California: Jossey-Bass.
- Livingston, D. W. (1995). For whom the bell curve toils. *The Alberta Journal of Educational Research*, 16 (3), 335-341.
- Luke, A., Land, R., Christie, P., & Kolatsis, A. (2002). *Standard Australian English and language Education for Aboriginal and Torres Strait islander Students*. Brisbane: Indigenous Education Consultative Board.
- Luke, A., Land, R., van Kraayenoord, C. and Elkins, J. (1997). *Report of an intrinsic critical appraisal of the Year 2 Net continua and associated teacher support materials undertaken for the Queensland School Curriculum Council*. Brisbane: The University of Queensland.
- Luke, A. and van Kraayenoord, C. (1998). Babies, bathwaters and benchmarks: Literacy assessment and curriculum reform. *Curriculum Perspectives*, 8 (3), 55-62.
- Jacoby, R. & Glauber, N. (1995). *The Bell Curve Debate*. New York: Times Books/Random House.
- Marks, G., & Ainley, J. (1997). *Reading Comprehension and Numeracy Among Junior Secondary School Students in Australia*. Hawthorn, Victoria: Australian Council for Education Research.

- Masters, G. N., & Mislevy, R. J. (1993). New Views of Student Learning: Implications for educational measurement. In N. Frederikson, R.J. Mislevy, & I.I. Bejar, (Eds). *Test Theory for a New Generation of Tests*. Hillsdale, New Jersey: Laurence Erlbaum Associates.
- Mayer, E. (1992). *Putting General Education to Work: The Key Competencies Report*. Melbourne: Mayer Committee, ACE.
- McElwain, D.W., & Kearney, G.E. (1970). *Queensland Test Handbook*. Melbourne: Australian Council for Educational Research
- McElwain, D.W. & Kearney, G.E. (1973). Intellectual Development. In G.E. Kearney, P.R. de Lacey & G.R. Davidson, (Eds.). *The Psychology of Aboriginal Australians*. Sydney: Wiley.
- McIntyre, L.A. (1976). An investigation of the effect of culture and urbanisation on three cognitive styles and their relationship to school performance. In G.E. Kearney & D.W. McElwain. (Eds.). *Aboriginal Cognition: Retrospect and prospect*. New Jersey: Humanities Press.
- McNaughton, S. (2000). *Meeting of Minds*. Wellington: Learning Media.
- Ministerial Council of Education, Employment, Training and Youth Affairs. (1997). *The National Report on Schooling*. Melbourne: MCEETYA Secretariat.
- Ministerial Council of Education, Employment, Training and Youth Affairs. (1999a). *The Adelaide Declaration on National Goals for Schooling in the Twenty-First Century*. Melbourne: MCEETYA Secretariat.
- Ministerial Council of Education, Employment, Training and Youth Affairs. (1999b). *The National Report on Schooling*. Melbourne: MCEETYA Secretariat.
- Ministerial Council of Education, Employment, Training and Youth Affairs. (2000). *Achieving Educational Equality for Australia's Aboriginal and Torres Strait Islander Peoples. Discussion Paper*. [On-Line]. Retrieved from the World Wide Web in September 2001 from www.curriculum.edu.au/mceetya.
- Nakata, M. (1997). *The Cultural Interface: An exploration of the intersection of Western knowledge systems and Torres Strait Islanders positions and expenses*. (Thesis) Townsville Queensland: JCU.
- Northern Territory Department of Education. (1999). *First Steps: External evaluation*. Darwin: Northern Territory Department of Education.
- Northern Territory Department of Education. (2000a). *First Steps across the Territory*. Darwin: Northern Territory Department of Education.
- Northern Territory Department of Education. (2000b). *Early intervention project: Final evaluation report*. Darwin: Northern Territory Department of Education.

- Ozolins, U. (1981). Victorian HSC examiners' reports: A study of cultural capital. IN H. Banister & L. Johnston. (Eds.). *Melbourne Working Papers*. Parksville, Victoria: Sociology Research Group in Cultural and Educational Studies.
- Porter, P., Rivzi, F., Knight, J. and Lingard, R. (1992). Competencies for a clever country: Building on a house of cards. *Ideas in Education*, pp. 9-14.
- Queensland Review of Queensland School Curriculum. (1994a). *Report of the Review of the School Curriculum: Shaping the Future*. 3 vols. Brisbane: Queensland Government Printer.
- Queensland Review of Queensland School Curriculum. (1994b). *Shaping the Future: Summary of recommendations*. Brisbane: Queensland Government Printer.
- Queensland School Curriculum Council. (1999). *Evaluation of 1998 Queensland Year 3 and 5 testing program*. Queensland: Queensland School Curriculum Council.
- Queensland School Curriculum Council. (2000). *2000 Queensland Years 5 and 7 Tests: Test Preparation Handbook*. Queensland: Queensland School Curriculum Council.
- Queensland School Curriculum Council. (2001). *2000 Queensland Years 3, 5 and 7 testing program. Report to the Minister of Education*. Queensland: Queensland School Curriculum Council.
- Rose, S., Kamin, L., & Lewantin, R. (1984). *Not in Our Genes: Biology, ideology and human nature*. Harmondsworth: Penguin.
- Ross, M. (1984). Intelligence testing in Australian Aboriginals. *Comparative Education*, 20 (3), 371-375.
- Sackett, P.R., & Wilk, S.L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist*, 49, 929-954.
- Scribner, S., & Cole, M. (1981). *The Psychology of Literacy*. Cambridge: Harvard University Press.
- Spiro, R.J., Bertram, C.B. & Brewer, W.F. (Eds.). (1980). *Theoretical Issues in Reading Comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence and education*. Hillsdale, New jersey: Laurence Erlbaum associates.
- State of Queensland. (1996) *The Year 2 Diagnostic Net Handbook for Schools*. Brisbane: QSCO.
- Stewart-Dore, N., Bartlett, L., Hallinan, P., Moreton, A., Robert, C. and Woodrow, C. (1999). *External Review of the Year 2 Diagnostic Net. Executive Summary of the report to the Queensland School Curriculum Council*. [On Line]. Retrieved

from the World Wide Web in September 2001 at
http://www.qscc.qld.edu.au/year_2_diagnostic_net/reports.html#external.

Strickland, B. (2000). Misassumptions, misadventures, and the misuse of psychology. *American Psychologist*, 55, (3), 331-338.

van Dijk, T.A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

van Kraayenoord, C., Luke, A., Elkins, J. and Land, R. (1999). The Queensland Year 2 Diagnostic Net: A critical commentary. *Journal of Research in Reading*, 22 (1), 95-105.

Willis, S. (1998). First do no harm: Accountability and the numeracy benchmarks. *Curriculum Perspectives*, 8 (3), 70-77.

APPENDIX 1:

CASE STUDIES OF EFFECTIVE PRACTICES

SUMMARY OF CASE STUDY THEMES

Four schools were visited as part of this research. The schools were selected to cover a range of systems and locations as well as different proportions of Indigenous students attending. This was to allow us to get a broad based understanding of the issues around assessment of Indigenous students as a base for other stages of the research. While all of the teachers, Principals and support personnel that we spoke to were committed to improving the standards of literacy and numeracy among the Indigenous students they were involved with, we found that at classroom, school and systems levels most were still grappling with organising effective informed assessment. The level of understanding that students' culture and background might be implicated in their ability to achieve equitable performance levels seemed to be variable. We contend that the issue of culture fair testing be developed as a key understanding of those personnel involved in the education of Indigenous students as an immediate priority.

What we saw in schools did not suggest that systems had been successful in their charge of providing organised, effective support structures for schools in relation to assessment in general, and certainly not in relation to the assessment of Indigenous students in particular. Quality practice seemed to be the result of individual expertise, and at best school level support. As schools with large Indigenous student populations, particularly those in rural or isolated communities, are also those who are most impacted upon by staff turnover, this reliance on individuals is not adequate.

It became evident in several contexts that the lack of systemic diagnostic and tracking information, not only impacted upon curriculum and pedagogy, but also enabled deficit assumptions about the performance patterns of Indigenous students to be generalised and to remain unquestioned.

In the absence of systemic level organisation, support and direction, the following seem to be issues that impacted on schools' ability to provide appropriate assessment techniques and procedures for their Indigenous students:

- school level organisation based on the support of strong leadership, highly trained staff, and clearly stated areas of responsibility for the assessment team members;
- an awareness of Indigenous culture and education as mainstream issues, rather than as separate programs;
- Indigenous staff, either professional or for whom the school sees it as a priority to train and invest in;

- a broad level, whole school approach to working collaboratively toward effective teaching and assessment of Indigenous students;
- staff awareness of the unique culture and language of Indigenous students and how this might impact on teaching, learning and assessing. This would seem to be just as important in school contexts with small numbers of Indigenous students as for those with high proportions of Indigenous students.

What follows are brief descriptions of the schools visited, their contexts and communities, and a highlighting of issues that seemed, to the visiting researchers at least, to be important issues impacting on the assessment practices they relied upon.

WATERDALE SCHOOL

The school/community in overview

Waterdale is one of two Torres Strait Islander communities on its small island location, and has a population of around 200 people. The community is located on the northeast coast of the island around a bay with a gradually sloping vista oriented towards the south-east sea breezes.

Waterdale was the site of an early Anglican mission, and the imposing whitewashed original church still dominates the view to the esplanade when approaching the community by road. The community consists of people from many different backgrounds – including South Sea Islanders, Tongans and locals originally from other islands in the Torres Strait. The Waterdale community council leases a sizeable parcel of land from the original owners. This progressive, entrepreneurial community has its own building blocks company and has renovated the esplanade, is building a new IBIS store and motel, and plans are afoot for development of several new recreational facilities.

While most local people are native speakers of KLY, Torres Strait Creole is the clear *lingua franca* for most social interaction that occurs in Waterdale - particularly among the younger generations. English, however, is the official language of schooling and employment and there are some tensions for community members committed to improving the quality of SAE at the school in the extensive use of TSC by young people.

The school is designated Band 6, with approximately 60 students being educated across pre-school to year 7. For the first time this year, the pre-school age children have been grouped with Year 1 students, to provide a seamless and individualized transition into school learning. Initially TSC is used to help in this adjustment to schooling, with SAE gradually becoming the dominant language of instruction. After completion of primary education, a small yearly cohort has to go away for secondary schooling, usually as boarding students at mainland schools. This is an occasion of considerable moment for both young people and their families, with the cost being a considerable burden on the finances of most.

The majority of the teachers, teacher aides and other school staff are local people, long-term island residents and experienced in the schooling system. At the time of visiting, the principal and three other staff were non-Indigenous.

Key issues for community engagement and sustained school response

Improving student skills of reading comprehension, research, library and ICT usage

The principal has assembled a considerable corpus of data on aspects of literacy and numeracy attained by Waterdale student over recent years. The school was the winner of the National Literacy Week Award for Queensland in 2000 because of the extent of student improvement demonstrated over the preceding three years on the Year 2 Net, and Years 3, 5 & 7 state-wide tests. However, it is apparent that these results were obtained with considerable scaffolding assistance from teachers and TAs during the testing period. The more recent results demonstrate a considerable variation in attained standards among students, and a lower overall average obtained under strict test conditions.

It is also apparent from the data, that a major issue is the need for more systematic development of higher-order skills and processes centred on reading comprehension and enhanced through research projects and computer usage. Current patterns of student reading and writing demonstrate a more limited and unsophisticated understanding of SAE structure and use in different contexts.

Two Indigenous staff have developed expertise in Productive Pedagogies and de Bono's Six Hat Thinking Skills program and are introducing other staff to key teaching/learning strategies designed to enhance student skill development. As well, a major investment and funding issue is anticipated in upgrading the school's currently limited capacity to provide sufficient opportunities for integrated use of computers for student research, word processing and literacy development. A set of grant applications has been submitted, that if successful should enhance the learning of school students, RATEP students and parents alike through provision of more computers.

Preparing students better for the transition to secondary schooling and SAE immersion

Students at the school enjoy patterns of adult support and levels of professional supervision that cannot usually be sustained by the secondary schools they will later attend. Specifically, the school has a designated intervention room and before-school intensive program for selected students, makes daily use of Indigenous Teacher Aides to scaffold student learning, and is about to embark on a ATAS homework centre initiative.

While these strategies should act to raise the overall average outcome standards for entry into secondary, a critical related issue is the need to develop greater student independence in their motivation and behaviours so that they are likely to adjust well and succeed reasonably when they begin secondary studies. Less reliance on one-on-one adult support, more involvement in small group research projects and taking greater individual responsibility are viewed as essential learnings across the Years 5-7 range. In addition, regular experiences in various 'test' environments are required to

familiarise students with a wider range of assessment instruments and methods and to be better prepared for the styles of thinking and analysis required to demonstrate successful performance.

Co-curricular activities such as leadership camps can assist in developing such transition skills, at least for some students. For the full range of students, it may also prove useful to establish sister-school relationships with key mainland schools that facilitate student exchanges and transitions and provide trial experiences of immersion in SAE and independent learning environments.

Developing skills of self-confidence, risk-taking and problem-solving

Many of the above initiatives also have the likely spin-off of increasing student capacities and emotional resilience in coping with change and developing their own ways of adjusting to new demands and challenges. In particular, there appears to be the need to develop an expectation that there are always options and choices when faced with difficulties and unfamiliar situations. In short, a key learning is that it is better to have tried and failed than to either give up or to always assume that adults will protect you from possible negative consequences of particular choices or actions taken.

Auditing and organizing available resources for teaching and learning

In common with most remote, Indigenous schools Waterdale has to regularly deal with sustainability issues associated with turn over of principals, teachers and TAs. Professional staff, in particular, tends to bring with them favourite strategies and resources that they consider more likely to generate greater impact or success. However, there are often unforeseen longer-term consequences of such earlier decisions about resource priorities that can seriously limit flexibility and integration in school responses.

On the other hand, if considerable care has been taken to audit all available resources and build selected strategies into a whole-school approach then it is more reasonable to expect that incoming principals and teachers will resist the immediate urge to change things and stay with the agreed directions until they have considered all the available evidence. In brief, it is not a shortage of useful ideas that may be limiting the potential achievements of Waterdale students but rather insufficient care taken to ensure sustainable implementation of curriculum initiatives and to succession plan for all key school staff positions.

Increasing teacher and parental expectations for high-quality student learning outcomes

Waterdale has a recent history of relatively high levels of commitment and involvement by key parents and community members, particularly women, in the school. This group is clearly also looking for ways to broaden and deepen the extent of parental participation in school activities, including more consistent monitoring of homework and reading. The current principal also has daily meetings with all school staff, and a separate one with teachers, to generate and maintain a broad-based commitment to the agreed change directions.

In particular, there appears to be a need to balance the emphasis on a support and intervention culture that scaffolds early student learning attempts with a subsequent

culture that encourages taking of risks in independent and cooperative actions and rewards more creative individual and small group responses to new situations.

As more secondary schools with boarding facilities set higher entry requirements it will become even more important for student to be able to set and attain progressively harder goals for themselves. Role models and mentoring may assist, but more resilience in the face of difficult challenges is perhaps a more important and portable capacity. High expectations, clear directions and sustained effort are all required by most students to succeed in further learning and employment contexts – particularly in mainland, SAE settings.

ST MARIE SCHOOL

The school in overview

St Marie School is a systemic urban Catholic primary school in the inner north east of Brisbane. Declining enrolments resulting from its location within an aging neighbourhood, and being surrounded by numerous other catholic schools eventually threatened to result in the school's closure in the late nineteen nineties. The school community and staff rallied to ensure this did not occur, and eventually an initiative put forward by the Parish Priest resulted in the school remaining open. In 1998 the school was designated as the School for Reconciliation for Education and Truth in History.

This designation has resulted in an increase in the proportion of Indigenous students attending the school, all of whom travel to school by bus from surrounding suburbs some of great distance from the school. It has also had implications for the school in relation to funding, with all staff agreeing the school is now well resourced.

The Principal has only been at the head of school administration since mid 2001, but has already begun initiatives to reform the school in the following areas:

- portfolio based assessment
- whole-school planning of curriculum units
- community and industry partnerships

The school is now firmly structured as a multi-age school, with students being grouped into preschool, years 1/2, years 2/4 and years 4/7 groupings. Teacher aides and other auxiliary staff regularly work within the classrooms resulting in low adult child ratios for much of the school day.

The student's in overview

St Marie has approximately 80 students, of which approximately 40% are Indigenous. More than 80% of the students were born in Australia, others were originally from countries such as Papua New Guinea, Indonesia and New Zealand. More than 70% of the student population is presently in the first three years of school, leading to the assumption that the school will continue to grow in the next few years. Unlike the locally living non-Indigenous students, the Indigenous students travel to school from a broad range of suburbs, being transported in by bus. This has implications for home school links, the students feeling part of a community, and parent input into school decision making.

Many of the Indigenous students have family connections to other areas of the state. The Principal estimated that approximately 20% of their student population was transient, with at least some of this number coming in and out of the school several times during their primary school career.

No school-based records of students' results in state-wide testing initiatives are readily available. However both teaching staff and support staff seemed to agree that the Indigenous students were performing below expectations, and were over represented in intervention and learning support groups.

The school operates on the assumption that the Indigenous children speak English as their first language. No Indigenous students have been assessed using the National ESL Bandscales. Several teachers and school officers expressed the opinion that the Indigenous students at the school were urban and that this limited the impact of Aboriginal English on their response to school. There was a sense that any language issues were not confined to the Indigenous students as portrayed in the following comments:

*I don't think really in here it is an issue. They're fairly urban.
It's not really Aboriginal English - we can all do different
languages can't we?*

(class teacher)

and

*some use Aboriginal English in school, but it isn't specific to
Indigenous kids anyway.*

(teacher aide)

The Principal and the Support Teacher Learning Difficulties affirmed the fact that Indigenous staff did not think that Aboriginal English was an issue for the Indigenous students at this school. A program set up in previous years to address children's use of Aboriginal English, had been perceived by several Indigenous staff as being deficit in formulation, and after a trial was not continued as a regular program within the school. The Home/School Liaison Officer did believe that there were issues in relation to written language that might cause problems for Indigenous students, although she did not elaborate on what these issues might be.

The staff in overview

The teaching staff consists of five classroom teachers to cover 4 groups, as well as specialist staff in the following areas; Health and Physical Education, Expressive Arts, Library, Learning Support, English Second Language. None of the teaching staff are Indigenous. There are several staff members who are new to the school in the 2002 school year. The school also has fourteen school officers, of whom five are Indigenous. These officers range in the number of hours that they work and in the tasks that they perform within the school, however the Indigenous officers tend to be involved in student-based work within the classrooms as the major part of their work responsibilities.

An integral member of staff is the Home / School Liaison Officer. In this role, she works closely with Indigenous families from the point of enrolment, meeting with all families once they identify themselves or their children as being Indigenous. Having been involved in the school for over 10 years, she now sees her role as being very much about making home visits and regularly meeting with Indigenous parents so that she can represent them and their views within the school community decision making process.

Several new staff highlighted the importance of the Home / School Liaison officer and other Indigenous School Officers, in their own development as effective teachers within the context of the school. One teacher went so far as to say that working closely with the Home / School Liaison Officer had allowed her to not only become more aware of the context and the students she was now working with, but also more aware of the fact that her own background was implicated in how she viewed students and events within her classroom.

The teacher librarian is also recognised by the staff as a resource in relation to Indigenous issues and students. Several teachers stated that the librarian was knowledgeable and was able to encourage an Indigenous cultural perspective throughout the curriculum. Part of this involved the organisation of visits from Indigenous guests and community members to coincide with curriculum units and school events. Hearing the voice of the Elders had been a program where Indigenous and non-Indigenous community members had visited the school to offer their expertise and knowledge to students and staff.

Broader political and social contexts

It is difficult to summarise the broader political and social contexts relevant to the Indigenous school population of St Marie. These children live in different locations, come from very different backgrounds, and belong to very different communities. Some have close ties to other Indigenous communities and spend time living in those communities. Community. Others may count their urban community as that which is most important to them - but it is unlikely that this urban community is related to the primarily non-Indigenous community surrounding the school. However as succinctly explained by one Indigenous staff member:

I don't know how they feel, but if they didn't feel it was working for them they wouldn't send their kids here would they?

Indigenous families are choosing to send their children by bus to this school rather than access school options closer to their homes. The reasons for this would seem to depend on individual family considerations.

Community input

Getting community involvement from parents remains an issue for the school in relation to both Indigenous and non-Indigenous families. There are no Indigenous parents who attend the P and F meetings regularly. Traditionally the school had serviced a local neighbourhood which seemed to be disenfranchised from the institution of schooling and thus had not been actively involved in the school. With a large proportion of students now living considerable distances away from the school, the more recent additions to the school community create other problematics in relation to inclusion in school activities.

Despite this the Home/School Liaison Officer reported that she believed that the Indigenous families did perceive the school to be a *friendly, family school*. She did not believe absenteeism to be a problem within the Indigenous student population, although she did recognise that pressures on families living in urban contexts without always having access to extended family or community support did have a major effect on student learning. While the Home/School Liaison Officer did identify that it was difficult to *get a lot of parents in*, she felt strongly that her role with its focus on home visits and support for families enabled the parents to feel that they had a voice within the school and were having their interests and needs put forward.

While the school does have an ASPA committee which meets regularly, attendance is usually limited to two Indigenous School Officers and two to four other parents. The Principal had organised for the bus to provide transport for parents to attend the night time meetings, but this had not increased attendance greatly. The committee and the Principal were looking for other solutions to increase membership. The Home/School Liaison Officer suggested that Indigenous parents were happy to let her know if they had something they wanted discussed at an ASPA meeting and thus did not feel the need to attend themselves.

Approach to organisation and curriculum

The school is organised with multi-age groupings in an attempt to demonstrate an awareness of the developmental needs of students. There was awareness that this structure also had other benefits for Indigenous students within the school:

Multi-age grouping works well with the Indigenous kids, because they have the same teachers. If the teachers hardly change they have the chance to get confident enough with them to talk.

(Home/School Liaison Officer)

In fact all staff members spoken to reported a belief that the multi-age grouping organisation was an effective way to approach teaching and learning within this school context. The Learning Support Teacher highlighted the link between this form of organisation and assessment when she stated:

because of the nature of multi aging, it is a requirement to acknowledge the amazing variability in levels. Because of the variable reading levels in rooms, teachers must be in control of assessment and intervention.

Approach to Indigenous education

Since being designated the School for Reconciliation for Education and Truth in History, leadership and staff at St Marie have found themselves contending with what that might mean in terms of curriculum and school initiatives. The new Principal suggested that finding ways to 'be' a model for Reconciliation for Education and Truth in History had meant he and the staff had had to seek community partnerships with other than education funding bodies. He reported that they were trying initiatives to see what made a difference, but were very much learning and piloting as they went along. It seemed that a particular difficulty had been to find ways to create a new curriculum for Reconciliation, which could become much more than a standard curriculum with add on programs for particular groups of Indigenous students.

Some students are involved in an Indigenous Dance Group, which staff felt had been an effective program to provide a cultural link between community and school. Dance had also been used as part of a behaviour modification program for senior boys within the school, with a visiting Indigenous dancer working with a group of students over time. This program had only been open to Indigenous students.

One initiative planned for term two is a whole school integrated curriculum unit, that will result in the creation of a Reconciliation / Peace Garden. The planned activities will involve Indigenous Artists, Health Workers and Community members in supporting students and teachers in the redevelopment of a section of the school grounds. It is envisaged that the curriculum unit will involve students in learning around the concepts of continuity and time, cultural identity, and space and time. This curriculum unit seems to have the potential to be an important program for both Indigenous and non-Indigenous students, and consequently might help the school to work toward including all students in the new curriculum for Reconciliation.

The school had also managed to attain funding for a training program in traditional Indigenous games (Choopadoo - NAIDOC week sport program. Edwards 1999). The training would mean students would become expert trainers and would then go out to other schools to teach students how to play selected games. As the other schools to be visited did not have large numbers of Indigenous students this program again has the potential to push the conceptualisation of a curriculum for Reconciliation beyond just relating to Indigenous students.

Stemming from a concern with nutrition amongst the Indigenous students, the Principal has also involved the school in a community based project funded by the Queensland State Health. The project is being organised by a nutrition reference

group and involves several local schools. The initiative will result in a fruit and vegetable co-op being set up within the school, and will enable families to access cheap quality fruit and vegetables. To build on the strengths of this project, the Principal has also organised to use a patch of adjacent land as a vegetable garden which will be managed by the upper school students within the school.

The school has also been a trial school in 2001/2002 for the ATAS in-school program (this replaces previous out of school tutoring arrangements). This has led to several existing Indigenous teacher aides receiving further inservice and employment hours, as well as the employment of several new Indigenous tutors from

Part of the reasoning behind increasing the preschool program to a full time five day a week unit - an initiative which occurred for the 2002 school year - had been providing high quality, full time early childhood education to Indigenous students. However it has proven to be an asset to non-Indigenous working parents who pass the school on the way to work in the city, more than Indigenous parents who would need to take advantage of the bus transport to access the preschool unit. Of the full preschool class only 2 students are Indigenous, and one of these students is a child of a member of staff at the school.

Approach to classroom learning programs - language and literacy education and numeracy education

The teachers at St Marie approach their planning as an individual activity. This is to be expected within a context where several teachers are new to the school. The administration of the school seems to have a notion that more collaborative planning will occur in the future as the staff begins to take responsibility for the direction of the school. However the importance of classroom teacher based responsibility for planning, teaching and assessing their students in the multi age setting was continually highlighted during discussions. This broadly spread leadership should be an effective foundation as the curriculum moves toward becoming one for Reconciliation in Education.

Several of the newly appointed teachers expressed how the context of teaching at St Marie was very different from what they had experienced in previous teaching positions. One teacher explained:

My views have changed in just seven weeks. I have to make it suit them, but the class is so different from what I'm used to.

Teachers reported planning using integrated units, with an emphasis on making the content relevant to students, and learning activities which allowed for different learning styles and different attention spans. One lower school classroom teacher had made *Jolly Phonics* a major part of her literacy program, believing that the kinæsthetic nature of the program was helping the students in her class with their early phonic learning.

A school based funding decision had been taken to provide access to a Speech Pathologist within the school. The private Speech Pathologist will work for 3 hours a week primarily with Indigenous students. This move seems to address a perceived

need for expert language instruction without entering the Standard Australian English / Aboriginal English debate that had caused tension in previous years between Indigenous teacher aides and non-Indigenous teaching staff.

There are several other school-based initiatives that approach intervention in literacy and language and in numeracy. These include access to large amounts of teacher aide time for in class work, support-a-reader, support-a-maths learner, a speech and language early intervention program and occasional withdrawal of students from class when deemed to be appropriate by the class teacher and the STLD. The strong foundation running under approaches taken to literacy and language teaching and intervention is that the classroom program should deliver for students rather than they being withdrawn to access supplementary programs.

Approach to assessment

The school staff is presently involved in working collaboratively to build an approach to assessment which works for students and staff in the school context. This approach is seen as an answer to the shortcomings that, particularly the Principal, see in the state-wide testing regime. Having been an outspoken critic of the benchmark and testing agenda for several years, the Principal is very conscious that critique without alternative solutions is not an option for systems and schools. To this end he is leading the staff toward setting up what he believes will be a viable alternative. Based in the tenets of portfolio assessment, the school-based approach is still in its infancy stage and because of this it is difficult to ascertain whether it will be a viable approach to accountability at the school.

While students were involved in the year 3, 5 and 7 Tests in Aspects of Literacy and Numeracy in 2001, under the present leadership, this decision was made due to perceived increasing system based pressure to be involved in the state-wide testing program. The Principal and several teachers did not believe that being involved in the testing program was implicated in parent choice to send their children to a systemic Catholic school. Parents were informed of the Testing program and when it would occur. When reports were returned to the school, the Principal used the newsletter to notify parents of their arrival, instructing those interested to make an appointment with him if they wished to be given details of their child's performance. The justification for this method of dissemination of results was that it allowed for the results and report format to be explained to parents. A very small number of parents chose to access the results. The 2002 teachers did not seem to have knowledge the 2001 results of their students from the testing program. Student results in the year 3, 5 and 7 Aspects of Literacy and Numeracy tests were thus perceived to have had a very insignificant position in the school's assessment approach.

The Principal's objections to the state-wide testing regime's approach to assessment included the incongruence between grade-defined standards and the philosophy behind multi-age education. He also expressed concern that the tests were incapable of linking into a learning cycle, with he believed no research available to demonstrate that the results from low stakes, norm referenced testing was used by teachers to enhance their teaching nor the learning of their students. In fact the Principal strongly disagreed with the linking of norm referenced testing and funding for essential intervention and support. He also believed that on a broader scale the testing regime

could lead to a testing culture emerging with the potential to define what education is about in our education systems, a dangerous potential if the form of the tests encourages conformity, compliance and control rather than celebrating difference and diversity (summarised from notes made by the Principal).

The portfolio assessment approach has been adopted as a means of complying with the shared beliefs of the staff on assessment and reporting. These shared beliefs are listed within the staff handbook as being:

- Assessment is not to prove but to improve
- Assessment should allow teachers to individualise and to meet the learning needs of each student
- Assessment should be seen in terms of individual progress
- Assessment methods should make students feel wanted in themselves and give them a sense of self worth
- Assessment methods should make students feel they have achieved something worthwhile
- Assessment methods should be clearly defined for all and open to processes of validity, giving students a sense of fair play
- Assessment methods should recognise individual differences, talents and learning styles and give students a sense of being accepted
- Assessment methods should make students feel they are worthwhile and give them a sense of confidence
- Assessment methods should allow parents to be reliably informed of their children's progress
- Assessment should be continual and ongoing
- Assessment should allow students to participate in assessing their own learning and act on the basis of that assessment.

Several staff members in the lower school and the Principal discussed their support of the Year 2 Diagnostic Net, particularly the developmental continua. This seemed to be because of the developmental nature of the continua, which fitted well with the beliefs of the staff about how students learn and also complimented the other assessment initiatives occurring in the school. One of the Year 2 teachers also explained that the validation process allowed her one on one time which she did not often have access to within the daily activity of teaching. The Year 2 Diagnostic Net continua are being included in the lower school portfolio framework, and are being considered - perhaps in a less detailed form - for upper school portfolios.

Funding and resource issues

All personnel who were asked clearly stated that the level of funding now available within this school context was high. This was a direct result of the school's designation as the school for Reconciliation in Education and Truth in History. It was generally perceived that this had had positive impacts on the programs and learning experiences supplied for students within the school. However the principal has also worked with staff and community to access other means of funding special programs. This is exemplified in the school's involvement in projects funded through Queensland Health, and the Department for Sport and Recreation.

GREENWOOD STATE SCHOOL

The school in overview

Greenwood State School is located in one of Education Queensland's northern districts and services the community Greenwood, a suburb of a large regional city. Despite some recent signs of change, this community has been recognised as being in a low socio-economic area with high unemployment. Recently the predominantly Aboriginal and Torres Strait Islander school population has seen a rise in the overall number of white Australian students, mostly of low socio-economic background attending. However 65% of the students are still of Aboriginal decent and 25% are of Torres Strait Islander background.

The school is committed to the community and to its place within this context. It has a special place in the Indigenous education in the state and was among the first provincial urban schools to employ Aboriginal and Torres Strait Islander teacher aids, support service and trained teaching staff. It has been a prototype site over the past two decades for research and development programs, for university-based intervention programs and for Aboriginal and Islander teacher education. It has also prototyped innovative programs that have addressed specific problems such as truancy, absenteeism and poor nutrition. For over a decade, it has built and maintained one of the strongest Indigenous cultural studies programs within the state system..

In recent years, the school has achieved marked success with its literacy curriculum initiatives and enjoys a productive and engaging relationship with the Greenwood and district Indigenous communities.. These positive outcomes may be attributable to the unique combination of components of community outreach and pedagogy that have been developed by the staff in response to the identified needs of the community.

Under its current leadership the Principal, the Literacy Coordinator and the Aboriginal and Torres Strait Islander student Coordinator work closely together. This has also involved direct engagement and contact with parents, with community Elders, and other key stakeholders. Everything in the leadership and running of this school reflects an abiding tenor of commitment to Indigenous education.

Greenwood State School receives additional funding through the Literacy Enhancement Special Program School's Scheme, the Indigenous Education Strategic Initiative project and the Community Access School's Project, which is a partnership agreement between Housing Queensland and Education Queensland.

The students in overview

A significant number of the Indigenous students attending the school either live in single parent families or their main caregivers are family members other than parents. The school's population is characterised by substantial seasonal movement of Indigenous children between Greenwood State School, Palm Island and Cape and Torres Strait schools. The raw mobility rate of students during 2000 was 83%.

Every year, Greenwood has a number of students (usually 20-25 in number) who have experienced many different schools and who as a result do not have basic reading skills. Sometimes these students only remain at Greenwood for a short period of time. The school has responded to the needs of these students with intensive programs, to improve reading accuracy and comprehension for students with very limited literacy skills.

Cross-agency referrals are used extensively at Greenwood, This has been largely at the initiative of the current Principal who has responded to the health concerns of the student body through the establishment of a Special Needs committee. Within the last two years, 25% of the Indigenous students have had hearing loss documented and a Soundfield Amplification System has been acquired for use in the school.

The staff in overview

A preschool teacher, 8 classroom teachers, a support teacher learning difficulties, a cultural studies teacher, a librarian and a Principal are employed at Greenwood State School. Years of service of this staff range from 2-25 years.

Greenwood is allocated 52 hours of General Teacher aide time per week, plus an additional 20 hours of Teacher Aide time per week due to the percentage of Aboriginal and Torres Strait Islander students attending the school. This amounts to ten Aboriginal and Torres Strait Islander Tutors at the school who offer direct assistance to classroom teachers.

Guidance Officer and Behaviour Management Support Teacher services are available on a regular basis. Two Home Liaison Officers are employed at the school to assist with student issues and a Community Development Worker is employed on a casual basis to encourage the Aboriginal and Torres Strait Islander Community to be more involved in the school and the students.

The focus for professional development and training in 2001 was to work toward implementing outcomes based education. Previously, Greenwood has invested substantial time and effort in a coordinated approach to literacy education, the trailing and prototyping of composite approaches (eg. Sound Way) and to phonemic awareness and direct instruction in the code. These investments appear to have yielded evidence of sustainable improvement.

Though there is not complete consensus on all aspects of its approach and operations, there is evidence at Greenwood that strong leadership and focus on the themes of literacy and Indigenous culture have drawn the staff's efforts together.

School Organisation

The Greenwood State Preschool is located off-campus some 3 kilometres from the Primary school. It is a 1-unit facility. Greenwood State School runs traditional year cohorts from Years 1-7. The school is currently trailing a program where teachers stay with the same group of students from year 5 to year 7. This is being trailed as a prototypical middle years and transitional strategy.

Community Input

Community liaising has virtually eliminated what was previously chronic absenteeism and truancy. The school maintains contact with parents and the community through a number of avenues including newsletters, assemblies, noticeboards, letters and meetings, the P & C, the ASSPA Committee and through media reports about the school. The community has many opportunities to be involved in the school and this is evident in the literacy plan outlined below. The current administration has been successful in making the school the responsibility of the whole community. One community elder commented that:

...it's sort of like the Musketeers – all for one and one for all
(Community Elder).

The suburb of Greenwood is currently being refurbished and revitalised by the State Government under the Urban Renewal Program. There is some evidence that the school leadership has been able to tap into cross-government and interagency support and cooperation with a focus on Indigenous education.

Approach to Indigenous Education

Cultural diversity is celebrated and respected within the Greenwood school community. Curriculum offerings ensure cultural diversity is strongly embedded within the school ethos. The Cultural Studies Program is a small, focused program tailored to exclusively meet the needs of the Greenwood community. Non-Indigenous students' participation is welcomed, encouraged and quite high. The tenor of the school is one of cooperation, where violence and bullying are not tolerated. Incidents around race are managed effectively and immediately.

In the last year, the literacy outcomes of Greenwood have surpassed expectations with Indigenous students' school reports recording significant improvement in all areas of literacy. By like school measures, its scores are above average.

Languages Profile

Many of the Indigenous students at Greenwood speak Aboriginal English or Creole as their first language. Making the transition to SAE and school in general, is taken by at least some of the staff to be a simpler task for particular groups of Indigenous students than other;

They (the children from the Torres Strait) don't seem to have any language barriers... they don't speak heavy Creoles. In Islander families, school is highly prized, many of the kids go on to be school captains
(Year 2 Teacher).

Despite the seasonal movement of children, the proximity to a large urban centre, media, and SAE peers, teachers and community members on a daily basis makes code-shifting common but on observation did not seem to be particularly visible classroom practice.

The positioning and use of ESL expertise

Greenwood State School has an active special needs committee that is spear headed by a Learning Support specialist. Eight of the teachers at the school have specific ESL education and there is active recruitment of Indigenous staff.

Language and Literacy Program

The staff has recognised that encouraging a book culture is important to the teaching of reading within the Greenwood context;

when I get my year 2's at the beginning of the year, they are about the level of mid year-one and many have not been to preschool'

(Year Two Teacher).

literacy does not have personal meaning for them. When they come, it's their first introduction to books

(Learning Support Teacher).

In response to these issues, Greenwood State School has established a highly innovative holistic language program, The Greenwood Literacy Program which aims to improve literacy outcomes for students, integrates three processes: intensive explicit instruction to develop basic coding skills; a community/cultural emphasis, and use of appropriate resources and information technology.

The rationale for the program derives from an investigation of national and school concerns regarding literacy outcomes. Several criteria have been used to track literacy outcomes, including standardised testing for reading and spelling, school records and anecdotal notes.

The needs of students are addressed within an awareness of the unique social and cultural environment of Greenwood;

engagement with literacy is what is needed...honour them in their knowledge and their culture

(Cultural Studies Teacher).

The literacy program involves a whole school approach but while the Special Needs Committee and the ASSPA Committee oversee the program, classroom teachers are responsible for its implementation.

I use a lot of scaffolding...hands on, try and discuss...lately I've been using big books followed by modelling and scribing...whatever works!

(Year 2 Teacher).

Although the language program sits neatly within a whole school plan, some concern was expressed that although the plan is prescriptive in nature, it is still open to interpretation. There was some tension over whether the program should remain

prescriptive and highly scaffolded, or whether it was time to move toward an approach that allowed for more teacher decision-making.

This tension between a heavily scaffolded program and classroom autonomy was evident in some of the discussions. But the point is that the school had developed a whole school program and appeared to have achieved results. Whether those results translated into a balanced literacy program across the board was an issue which would require further investigation, particularly given the state-wide and international data that indicated potential 'washout' and 'wastage' effects from intensive and strongly focused early intervention studies. Longitudinal studies of development of comprehension, writing and critical literacy would be needed. These would prove difficult given the confounding variables of high student mobility and staff turnover.

Curriculum Initiatives

The Greenwood Indigenous Literacy Program is part of the Greenwood Literacy Improvement Program. The program targets all Indigenous students in years 1-7 and was developed in response to low Indigenous student attainment in literacy - especially in comparison to like-schools - from data from the Year 2 Diagnostic Net, the years 5 and 6 tests and school based student reports. The initiatives are structured to cater for the unique composition and challenges of students at the school.

The initiatives fall within three broad strategies:

Improving basic literacy skills

Tutoring within an in-school homework program staffed by skilled Indigenous tutors, and a Reading Improvement Program based on Reading Recovery guidelines and staffed by highly trained teacher aides.

The Sound Way: A multi sensory, multi modal program aimed at all year levels. Some areas of focus include phonics, punctuation and handwriting.

Phonemic Awareness: 27 identified students receive individual phonemic awareness instruction from a speech-language pathologist on a weekly basis.

Reading Recovery: Delivered by a highly trained and stable staff member, who has the knowledge and experience to adjust the program's strict procedures.

A focus on community involvement

Involving parental tutors, parental engagement programs and teacher aid employment.

Cultural Studies Program

A Literacy skills program taught within a Cultural Studies framework.

This rigorous focus on basic skills sat within a longstanding ATSI cultural studies program that used Indigenous volunteers to teach aspects of cultural knowledge and to assist and support readers. Students, parent and Elders were involved in developing high quality ATSI community studies models. Community projects including a

school-published newspaper, the Country Women's Association International Country and Study Competition, and extremely high-profile NAIDOC week activities. The Concentrated Language Encounters model was used in alignment with integrated units of study and community excursions, under the assumption that the CLE model would enhance students' writing, mastery of genre and engagement with functional and critical literacy.

NORTH HILL SCHOOL

The school in overview

The North Hill School is a medium-sized P-7 Catholic school (352 students), situated in a Northern Urban suburb. The school has been open in its present form since the late eighties. It was formed through the amalgamation of four local parish schools.

The school has recently restructured and now structures its classes as multi-age units. There are presently 7 multi-age classrooms. The school now consists of a Junior School (Year 1-3), a Middle School (Year 4-5) and a Senior School (Year 6-7). The restructuring is partly a response to state-wide testing results: benchmark tests and scores did not match Queensland like-school norms in 1998 or 1999.

Along with the move to multi-age classrooms, the school has developed a school-based integrated curriculum and reporting system, begun in 1995. This curriculum has been refined over successive years, with recent emphasis upon outcomes based curriculum planning.

The school community in overview

Teachers at the school describe it as:

a small, tight-knit school community, with clearly articulated goals of providing a caring and nurturing environment for individual students and families.

(Year 2 teacher)

Students

Of the 352 students, 12 claim Indigenous status. There have been up to 25 Indigenous students at the school in the past. Many of these Indigenous children are from families with long affiliation with North Hill School.

Despite the fact that the school is a fee paying school, only 51% of families pay full fees. According to the 2000 school annual report, almost one child in four attending the school lives below the poverty line.

Students live in various areas of the city, but approximately 40% live in nearby suburbs. As the school is close to a Defence Force facility, North Hills is a Defence Force Priority School. Children of Defence Force personnel do represent a proportion of the transient student population, however recent changes to Defence Force policy about personnel movement has meant that these families are now more geographically stable than in previous years (Principal). However troop movements often mean that school students must deal with absent parents for significant amounts of time.

Within the small Indigenous student body, there is considerable economic, cultural and linguistic diversity, however they tend to be stable school enrolments. Some families who are linked with powerful Indigenous advocacy groups, may be employed by these groups, and perceive themselves as fluent Standard Australian English speakers. This group appears to both value a traditional western education, and want to maintain links with their Indigenous culture. Other Indigenous families are less affluent, have close links with their Indigenous community, but seem to have more difficulty engaging with the school culture and hence negotiating for their child within that culture. A third group is those children of Indigenous descent who live with a non-Indigenous parent, and are not encouraged to identify as Indigenous. While teachers made few references to linguistic diversity within the Indigenous group, several references were made to cultural differences in learning style, particularly to students' comfort with asking for help, and preference for group or one-to-one teaching.

What is of interest in this case study is the degree to which Indigenous students who are a small cohort in a mainstream school are 'integrated' and provided with specific diagnostics, intervention and assessment.

North Hill School has a part-time teaching principal, 14 full-time teachers, 5 part-time teachers and 4 teacher aides (one full-time in the pre-school). There is also a part-time school counsellor, a part-time Assistant to the Principal - religious education, 2 administration staff and 2 Grounds staff. None of these staff members is Indigenous, however two of the teaching staff have had extensive experience with Indigenous communities, e.g. Palm Island.

Parent / community input

For Indigenous students, the main parent/community input is through the ASSPA (Aboriginal Student Support and Parent Awareness) Committee. At North Hill, this committee consisted of approximately 5 parents. The major emphasis of the group is cultural, with both Aboriginal and Torres Strait Islander dance troupes operating within the school. These troupes have implicit and explicit functions. An explicit function is to share Aboriginal & Torres Strait culture with the wider school community. An implicit function is to involve students who are not encouraged to identify as Indigenous in culturally important activities.

The ASSPA committee aims to encourage broader Indigenous involvement in the school, and possibly to broaden its advocacy role within the school community.

Approach to Indigenous education curriculum structure

Teachers previously have had access to inservice programs on Indigenous issues and reconciliation. The school administration and leadership had taken a visible leadership role in highlighting Indigenous issues at this school.

All teachers were in-serviced in First Steps in 1997. This in-service was then linked with curriculum development. After development of the curriculum guidelines in 1997, the school decided to provide more integrated curricula, and to reduce the reporting process. Those Indigenous students who required learning support, receive the support from the Learning Support Team within the school. The ASSPA Committee manages IESIP funding, which is allocated by the diocese office. Indigenous community input into curriculum delivery appears limited to Support-a-Reader tutoring and funding for Introductions to Computers, via IESIP funding. As well as this, IESIP funds the small scale Indigenous Literacy/ Numeracy Support Program.

The policy and program context for accountability and monitoring

Although it appears that the School provided good infrastructure and pastoral care for its small Indigenous cohort, the actual systematic use of assessment and reporting strategies was reliant on the expertise of the learning support teacher. The choice had been made to use Neale reading inventories, on advice from the diocese office. Each child had been periodically checked on a reading inventory. Because there were no system or state-wide norms, the actual comparative tracking of student data appeared to be difficult.

Statewide Testing

Full participation in the state-wide testing is not compulsory for schools within this educational system. North Hill School participated in state-wide testing in 1998 and 1999. This year the school completed the Validation Tasks and was moving towards full Net implementation.

When state-wide testing was carried out at the school, teachers report that some students found the experience difficult: some items were not related to the Queensland Curriculum, and students became quite anxious. Reactions to the state-wide testing process were not specific to Indigenous students: bright students managed well, but students with learning difficulties found the process overwhelming. For Indigenous students, the visual nature of some of the test items were seen as an advantage, but teachers commented that many of the perceived strengths of Indigenous students (e.g. multi-intelligences) were not drawn upon.

One important factor for the school was that Indigenous parents seldom came forward to discuss results of state-wide testing:

I couldn't understand the results for J (Year 3), and I never saw anything for (Year 6). I think B might have done something this year.

(Parent of two Indigenous students)

Perhaps as a result of this limited participation, state-wide testing does not seem to be closely tied to other accountability measures within the school.

Approach to Accountability issues with Indigenous students

Accountability to parents/ carers

This level of accountability is met through regular Parent Information Sessions (PINS), exposure of curriculum units through school and class newsletters, office and classroom displays, homework, and the school-based reporting instrument, the Brag Book. The brag book is a qualitative reporting mechanism, based on the Key Learning Areas as they relate to a particular theme. Reporting is colour coded, and consists of both teacher report and work samples that are selected from class work. Each child also completes a self-assessment, based on life-long learning skills. At the end of each theme-based teaching period, parents are asked to meet with teachers and discuss their child's learning outcomes.

Indigenous families are believed to prefer this more integrated reporting:

there is less paperwork, more visual recording, and the child can 'talk the parents through' the report relatively easily.

(Curriculum Coordinator)

The Brag Book is:

more visual, easier to understand.

(Indigenous parent).

Some families had described the earlier reporting format as a "*sea of ticks*" (curriculum coordinator) that was very difficult to understand. Class teachers also report that Indigenous families are more likely to come to the school for these regular reporting nights.

Accountability to the Indigenous community

The local CEO employs an Indigenous Literacy Project Officer (ILPO) to monitor Indigenous students' progress and advocate for these students in Catholic schools. There are some 400 Indigenous students in the local schools. The ILPO tests all Indigenous students on standardised literacy tests (Neale Analysis of Reading Ability – Third Edition, Waddington Spelling) plus qualitative assessment of literacy (Waddington Diagnostic Reading Test), numeracy (Diagnostic Mathematics Test – adapted for Queensland curricula, classroom observations and analysis of work) and classroom behaviour.

Attitudes to testing, for Indigenous students within the district system seem to be somewhat ambivalent. On one hand, there are criticisms of the inappropriateness of most tests in cultural and curriculum terms. On the other hand, there is acknowledgment of the need to monitor student progress, and track outcomes from funded programs for Indigenous students.

It may not be culturally appropriate, but if we don't do it, we don't know who to intervene with.

(ILPO).

Curriculum Accountability and Monitoring

As the major reporting mechanism in the school, the Brag Book, is largely qualitative, one perceived gap in program accountability and monitoring is a lack of appropriate quantitative measures of learning outcomes. One staff member was concerned that some teachers were using group administration of standardised word recognition tests to 'establish reading levels' for some class groups.

Funding and resource issues

Funding for Indigenous students comes from recurrent funds, Special Needs funding, ASSPA and IESIP. At the school level, apart from specified programs such as ASSPA, funding is applied for through the local CEO for specific cases of need, or specific projects.

The future

It is likely that enrolments at the school will increase in the next 2-3 years as more young families move into the area. The reputation of the School as having high levels of pastoral care for Indigenous families, and the receptiveness of the school leadership to Indigenous issues, were strong.

Specific recommendations

According to the ILPO, the effect of state-wide testing on some Indigenous school populations can be destructive to students' self-esteem as learners. It was reported how at one local school, Indigenous students became very disturbed during the Year 5 test:

One girl screamed. Having 'mental arithmetic' – as it was first, it blew them away.

(ILPO)

The context of test administration was reported to have some negative effects.

North Hill School is an interesting case, both to track those Indigenous children who are integrated and potentially invisible in small cohorts in urban settings and also as a case of local assessment and reporting systems at work.

It is worth noting that the state-wide tests, their administration and data reporting and analysis, are clearly not fully integrated into and accepted within the culture of this school and other Diocese Catholic Schools at least in this region. The Net was only in its earliest stages of implementation. While the school had put in place some diagnostic systems – some of which including the Brag Book were quite laudable and effective in engaging parents and students in portfolio-based evaluation – the staff had

had to largely rely on advice from the Catholic Education Office as to which diagnostic instruments to use. The result was a somewhat ad hoc mix of standardised tests, reading inventories and teacher-based assessments. There was little evidence of the shared vocabulary or professional development effects of the Year Two Net cited earlier in this report, or of the use of testing data to set school targets or benchmarks. In sum, it was difficult to get an overall picture of literacy and numeracy achievement with any degree of comparability or systematicity.
