
Copyright 2010 Annette Woods and Adele Amorsen
EVALUATION OF THE YEAR 1 LITERACY AND NUMERACY CHECKPOINTS ASSESSMENTS TRIAL – 2010

Annette Woods

and

Adele Amorsen

School of Early Childhood
Faculty of Education
Queensland University of Technology

Research Report Prepared for the Queensland Studies Authority
The Evaluation team would like to thank the school-based staff who have responded to the Evaluation in positive and professional ways. We would also like to thank the members of the Working Party and particularly the QSA Implementation team, who have gone well beyond what could usually be expected to support the Evaluation. The collegial approach taken by all involved has enabled the Evaluation to achieve considerable results in very tight timelines.

We would also like to thank the literacy and numeracy researchers and teachers who provided expert comment on the materials. Their insights were invaluable to this research.

We particularly thank Karen Dooley (QUT) and Beryl Exley (QUT) for their useful and expert reviews of the tools. The analyses influenced the direction of the analysis presented in Chapter 4 within this report. Their input was invaluable and is acknowledged here.

We thank Chrystal Whiteford for her expert work in presenting data and the statistical analysis presented in this report.
Executive Summary

Summary
The draft Year 1 Literacy and Numeracy Checkpoints Assessments were in open and supported trial during Semester 2, 2010. The purpose of these trials was to evaluate the Year 1 Literacy and Numeracy Checkpoints Assessments (hereafter the Year 1 Checkpoints) that were designed in 2009 as a way to incorporate the use of the Year 1 Literacy and Numeracy Indicators as formative assessment in Year 1 in Queensland Schools. In these trials there were no mandated reporting requirements. The processes of assessment were related to future teaching decisions. As such the trials were trials of materials and the processes of using those materials to assess students, plan and teach in year 1 classrooms.

In their current form the Year 1 Checkpoints provide assessment resources for teachers to use in February, June and October. They aim to support teachers in monitoring children’s progress and making judgments about their achievement of the targeted P-3 Literacy and Numeracy Indicators by the end of Year 1 (Queensland Studies Authority, 2010 p. 1). The Year 1 Checkpoints include support materials for teachers and administrators, an introductory statement on assessment, work samples, and a Data Analysis Assessment Record (DAAR) to record student performance. The Supported Trial participants were also supported with face-to-face and on-line training sessions, involvement in a moderation process after the October Assessments, opportunities to participate in discussion forums as well as additional readings and materials.

The assessment resources aim to use effective early years assessment practices in that the evidence is gathered from hands-on teaching and learning experiences, rather than more formal assessment methods. They are based in a model of assessment for learning, and aim to support teachers in the “on-going process of determining future learning directions” (Queensland Studies Authority, 2010 p. 1) for all students. Their aim is to focus teachers on interpreting and analysing evidence to make informed judgments about the achievement of all students, as a way to support subsequent planning for learning and teaching.

The Evaluation of the Year 1 Literacy and Numeracy Checkpoints Assessments Supported Trial (hereafter the Evaluation) aimed to gather information about the appropriateness, effectiveness and utility of the Year 1 Checkpoints Assessments from early years’ teachers and leaders in up to one hundred Education Queensland schools who had volunteered to be part of the Supported Trial. These sample schools represent schools across a variety of Education Queensland regions and include schools with:

- A high Indigenous student population;
- Urban, rural and remote school locations;
- Single and multi-age early phase classes;
- A high proportion of students from low SES backgrounds.

The purpose of the Evaluation was to:

Evaluate the materials and report on the views of school-based staff involved in the trial on the process, materials, and assessment practices utilised.

The Evaluation has reviewed the materials, and used surveys, interviews, and observations of processes and procedures to collect relevant data to help present an informed opinion on the Year 1 Checkpoints as assessment for the early years of schooling. Student work samples and teacher planning and assessment documents were also collected. *The evaluation has not evaluated the Year 1 Checkpoints in any other capacity than as a resource for Year 1 teachers and relevant support staff.*

**The Design of the Project**

The Evaluation has collected both qualitative and quantitative data through on-line surveys, interviews, observations. This data has been supported by the collection of student work samples and records of teacher planning and assessment processes. The analysis has taken a mixed methods approach to ensure that concepts can be discussed across the full data set, with the ‘particular’ of local and contextual factors also becoming visible.

The truncated time schedule of the Supported Trial has had implications for both those involved in the Supported Trial and for the Evaluation team. However, as importantly for the Evaluation, the actual timing of the Trial (August to December) has impacted upon the data collection processes. For example, while teachers and leaders have been very generous with their time, it was difficult to arrange visits for interviews in the last few weeks of the school year. The business of schools at this time of year has also been reported to us as the reason that some teachers did not take up the opportunity to participate in the on-line surveys. Response rates were adequate, but were not large enough to allow technical analysis of data. We originally planned to follow up on trends with selected teachers in the early weeks of the new year, however school and university disruptions caused as a result of natural disasters in Queensland also made this impossible. The quantitative and qualitative data collected from survey respondents has been supplemented with qualitative data provided by teachers and administrators in other forms and the design has been flexible so as to work around these issues. Data collection has been appropriate for the claims made within this report.

The study has involved seven major steps, conducted between September and December, 2010:
1. A review of literature on assessment approaches, moderation, early years assessment, and summative, formative and diagnostic assessment. This review was used as a template for investigating the use of materials based in a model of assessment for learning in the early phase of schooling.

2. Collaborative work with the Year 1 Literacy and Numeracy Checkpoints Assessments Working party throughout the duration of the project. This has allowed the documenting of the development processes of the Year 1 Checkpoints.

3. Desk analysis of approaches to early years’ assessment taken by other systems across Australia, and revisiting the past work conducted on Qld’s current approach to early years diagnostic assessment. This was used to allow investigation of the Qld approach within the context of systemic approaches more broadly.

4. Taking advice from expert literacy researchers and practitioners in the review of the materials against criteria set up from the literature.

5. Attendance and observation at trial events and training opportunities. This provided the opportunity to collect data and receive feedback from participants in the trial.

6. Surveys of: 1) participants at Moderation Sessions; 2) year 1 teachers (focus on June and October Assessments); 3) Prep teachers (focus on February Assessments); and 4) Facilitators of the Supported Trial process. This was used to collect both quantitative and qualitative data on the materials and processes of the Checkpoints from the perspective of those involved in the Trial.

7. Interviews of a select group of teachers after completion of surveys. This was used to sample local school and classroom use of the Year 1 Checkpoints.

Findings
F1. That all Australian State/Territory systems are currently engaging their early years’ teachers and students in some form of assessment in literacy and numeracy. These tools vary in content, emphasis, theoretical foundations and quality. The expectations of teachers and their reporting mandate also vary. The foundation of Qld’s current
approach is the Year 2 Diagnostic Net, which has been in service for sixteen years, despite numerous calls for its renewal or replacement.

F2. That the P-3 Indicators state expected learning for each year level and thus differ from taxonomies of literacy and numeracy development such as the Year 2 Diagnostic net continua. This difference is related to framing much more than it is to content.

F3. That there is evidence that many teachers had not used the Indicators as a framework prior to being involved in the Supported Trial. Some teachers had difficulty understanding the format, content, purposes and intent of the Indicators.

F4. That there is evidence that the Year 1 Checkpoints were generally well received by many teachers and school-based staff as items of assessment for the early years. This is not to say that teachers have not requested and demanded many modifications. These modifications were based on: the call for correction of errors; difficulties with understanding the purposes of individual assessment requirements; difficulties with the format or structure of specific assessments (particularly issues with a bias toward one-to-one individual assessment); issues related to making amendments to the assessments for specific students or cohorts of students; issues related to the quantity of assessment items and issues of time and workload; and issues of timing, time and workload more generally.

F5. That the additional support materials were generally well received by teachers. That interactive or digital versions of some of these support materials – particularly the DAAR – would support teachers in their use of these materials.

F6. That the expertise and professionalism of the QSA team supported teachers and administrators in their assessment work in schools. That there needs to be a variety of approaches to training and advice giving so as to accommodate teachers in a variety of contexts and the local constraints of these contexts (such as connectivity issues), and a variety of levels of teacher understanding, experience and IT capabilities.

F7. That, contingent on the suggested changes from chapter 4 of this report and the achievement of F8 and F9 below, the Year 1 Checkpoint Assessments generally follow the principles of quality early years assessment for learning. As such the materials could be used to support year 1 assessment and subsequent planning and teaching. They have not been evaluated in any other capacity.

F8. Issues related to cultural fairness of the Year 1 Checkpoints have not been adequately addressed to date.
F9. Issues related to the capacity of the Year 1 Checkpoints to deal productively with diversity of experience and specific needs of students, and issues related to building teacher capacity in differentiation of assessment items to ensure the processes of assessment are based in equity and justice have not been adequately addressed to date. Teachers reported making amendments to the assessment processes to accommodate the needs of a variety of students, but there was no advice provided about such differentiation.

F10. Early years teachers in Queensland have focused on validating children within the Year 2 Diagnostic Net processes. Despite whether this was the original intent of the Year 2 Diagnostic Net materials, there is much evidence to support that this is the case. Linking the Net validation tasks with identification of individual students for funding purposes has ensured that the assessment items be understood in this way. The Year 1 Checkpoint materials and moderation processes have the potential to refocus teachers toward using checkpoints to validate teacher judgments rather than validate children. This has the potential to focus assessment on future teaching and learning.

F11. That the social moderation process implemented as part of the Supported Trial supported consistent teacher judgments of students’ work in Year 1. As a professional development opportunity teachers rated the moderation process as useful. Moderation has the potential to support a focus on validating teacher judgments.

F12. That the Year 1 Checkpoint Assessments currently provide a framework and a suite of assessment instruments for use in monitoring students’ achievement in literacy and numeracy in Year 1. There are issues that need to be addressed with individual Assessments and with the package as a whole (see chapter 4). The Year 1 Checkpoints can currently be used at three points across the year to monitor students’ achievement toward end of year indicators in literacy and numeracy so as to inform future teaching and learning. They do not currently feature expectations, standards, and resourcing decisions that would allow them to be deployed beyond this purpose. Their link to an end of year expectation has implications for their deployment as a diagnostic instrument to identify students for intervention and/or funding.

Recommendations

R1. That Queensland systems work toward a consistent, quality and rigorous system-based approach to early years assessment and diagnosis of literacy and numeracy learning needs across P-3 as a matter of urgency. In the context of education reform in Australia currently, it is important that Queensland have a current and well-resourced, high quality, high equity approach to assessment in literacy
and numeracy in the early years of schooling. These decisions must take account of the measures already in place and their utility, as well as teacher workload, and the consequent delivery standards required to support a cohesive approach across P-3.

R2. That the QSA and/or EQ establish a working group to oversee the progress of a system-based approach to early years’ assessment in literacy and numeracy. The task of this working group would be to design and develop a coherent approach to early years’ assessment for Queensland schools. The group would require technical expertise in measurement, evaluation and assessment, along with expertise in current approaches to literacy and numeracy, Indigenous education, and early childhood education. Their purpose would be to achieve a state-wide approach to early years’ assessment that is coherent, culturally fair, rigorous, valid and technically advanced, as well as realistic in terms of teacher workload and the delivery standards required for implementation.

R3. That any decisions about the uses and purposes of the Year 1 Checkpoints be made in light of the findings of past reviews of diagnostic assessment in Queensland. That the calls for culture fairness, validity, reliability and equity be considered as paramount as assessment instruments and frameworks are put to work as part of a system-based approach to assessment of literacy and numeracy in the early years.

R4. That a review of the content of the Literacy and Numeracy Indicators as they relate to current understandings of literacy and numeracy, and current and future curriculum expectations of literacy and numeracy across the curriculum be conducted.

R5. That appropriate use of language to describe the Indicators as expectations of learning and not a map of student development be used. That relevant teachers and school administrators receive advice and training which includes opportunities to develop an understanding of the underlying theoretical foundations of the P-3(9) Indicators as a frame for assessment purposes.

R6. That if the Year 1 Checkpoints are to be used in Queensland schools, that relevant teachers and school administrators receive advice and training on the Year 1 Checkpoints, the principles of literacy and numeracy teaching underlying the materials and process, the principles of assessment for learning, achieving consistent teacher judgments, the use of assessment for future planning and teaching, and processes of moderation for validating teacher judgments.

R7. That if the Year 1 Checkpoints are to be used in Queensland schools, that the introductory statement of the Year 1 Checkpoints be extended to address not just principles of assessment, but also the theoretical
foundations of literacy and numeracy being deployed in the Year 1 Checkpoints, and advice on how the Indicators and Checkpoints are linked.

R8. That if the Year 1 Checkpoints are to be used in Queensland schools, that the advice given in Chapter 4 of this report be used to guide modifications. And that the QSA consider ways to simplify the assessment items as they are modified and alternatives are designed. This does not suggest a lowering of standards, but rather a consideration of issues of classroom implementation and organisation, teacher and student assessment workload and time, content and construct validity, issues of language and content, and the dynamics of individual (one-to-one) and group assessment in the early years of schooling.

R9. That if the Year 1 Checkpoints are to be used in Queensland schools, that the QSA and EQ develop and trial alternative instruments as they modify and further develop the suite of materials so that teachers and schools eventually be provided with a bank of assessments from which they are able to select instruments that work within their site-specific contexts.

R10. That this development process (R9) include the development and trial of culturally relevant materials for Indigenous students across a variety of contexts. That Indigenous expert educators, with a broad understanding of the particularities of location evident for some students in Queensland schools and technical assessment and literacy and numeracy expertise drive this design and development process.

R11. That teachers in the early years receive advice and training in approaches to assessment that are culturally fair and able to deal with diversity and specific student learning needs. That this advice and training also focus on how best to record, monitor, analyse and report the differences that students bring to assessment tasks in educationally useful ways.

R12. That a process of social moderation be implemented as part of any implementation of the Year 1 Checkpoints.

R13. That issues related to delivery standards, such as required funding for the assessment process, be considered as part of any implementation plans of the Year 1 Checkpoints.

R14. That positioning of the Year 1 Checkpoints in relation to the P-3 years of schooling be considered as part of any implementation plans of the Year 1 Checkpoints.
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY** .................................................................................................................................................. 3

**SUMMARY** ........................................................................................................................................................................ 3

**THE DESIGN OF THE PROJECT** ........................................................................................................................................ 4

**FINDINGS** ............................................................................................................................................................................. 5

**RECOMMENDATIONS** .......................................................................................................................................................... 7

**1. INTRODUCTION** ............................................................................................................................................................ 12

**THE PROBLEM** .................................................................................................................................................................... 16

**THE PURPOSES OF EARLY YEARS ASSESSMENT AND MODERATION IN A STANDARDS BASED REFORM CONTEXT** ................................................................................................................................. 18

  **SUMMATIVE ASSESSMENT** .................................................................................................................................................. 19

  **FORMATIVE ASSESSMENT** .................................................................................................................................................. 20

  **DIAGNOSTIC ASSESSMENT AND DIAGNOSTIC ASSESSMENT TOOLS** ........................................................................ 22

**2. CONSIDERATIONS FOR ASSESSMENT** .......................................................................................................................... 24

**INTRODUCTION** ...................................................................................................................................................................... 24

  **ASSESSMENT OF LEARNING** ........................................................................................................................................ 25

  **ASSESSMENT FOR LEARNING** ........................................................................................................................................ 26

  **ASSESSMENT AS LEARNING** ............................................................................................................................................ 28

**3. EARLY YEARS ASSESSMENT - POLICY AND PRACTICE IN AUSTRALIA** .............................................................................. 30

**INTRODUCTION** ...................................................................................................................................................................... 30

**EARLY YEARS ASSESSMENT** .................................................................................................................................................... 31

**THE AUSTRALIAN CONTEXT:** .................................................................................................................................................. 32

  **NORTHERN TERRITORY** .................................................................................................................................................... 34

  **VICTORIA** ............................................................................................................................................................................ 36

  **WESTERN AUSTRALIA** ......................................................................................................................................................... 37

  **AUSTRALIAN CAPITAL TERRITORY** .................................................................................................................................. 38

  **TASMANIA** .......................................................................................................................................................................... 40

  **SOUTH AUSTRALIA** ............................................................................................................................................................ 42

  **NEW SOUTH WALES** ......................................................................................................................................................... 44

**THE QUEENSLAND APPROACH** ........................................................................................................................................... 45

  **PREVIOUS EVALUATIONS OF THE YEAR 2 DIAGNOSTIC NET MATERIALS AND PROCESSES** .................................... 47

**CONCLUSION** ............................................................................................................................................................................ 49

**4. THE YEAR 1 CHECKPOINTS** .............................................................................................................................................. 51

**INTRODUCTION** ........................................................................................................................................................................ 51

**YEAR 1 CHECKPOINTS ASSESSMENTS TRIAL PROCESSES** .............................................................................................. 52

**ADJUNCT SUPPORTS PROVIDED TO SUPPORTED TRIAL PARTICIPANTS** ......................................................................... 53

  **ON-LINE TRAINING SESSIONS** ................................................................................................................................... 54

  **ONGOING ON-LINE DISCUSSION FORUM** ........................................................................................................................ 56

  **QSA STAFF SUPPORT** ......................................................................................................................................................... 58
LIST OF TABLES

**Table 3.1:** Name of the first year of schooling and the required starting age for each Australian State and Territory

**Table 3.2:** Summary of assessment items used in each State and Territory of Australia

**Table 4.1:** Number of schools in the Supported Trial according to relevant school categories.

**Table 4.2:** Numbers and percentages of participants who participated in the *On-line Training Sessions* and reported finding the sessions extremely useful, middle of the scale and not at all useful.

**Table 4.3:** Numbers and percentages of participants who participated in the *On-line Discussion Forums* and reported finding the sessions extremely useful, middle of the scale and not at all useful.

**Table 4.4:** Numbers and percentages of participants who contacted the QSA for advice or support and reported finding the contact extremely useful, middle of the scale and not at all useful.

**Table 4.5:** Numbers and percentages of participants who contacted other EQ staff for advice or support and reported finding the contact extremely useful, middle of the scale and not at all useful.

**Table 4.6:** Initial advice provided to Supported Trial schools on expectations and requirements related to Supported Trial participation.

**Table 4.7:** Number of respondents to each of three surveys and the school based roles of these respondents.

**Table 4.8:** Distribution of respondents across priority school categories n (number of participants in sample in each school category).

**Table 4.9:** Distribution of Responses across the 7-point Likert scale on level of agreement to statements about the materials generally. (Shading represents the most popular response for each statement).

**Table 4.10:** Distribution of responses across a 7 point Likert scale when year 1 teachers were asked to comment on their level of agreement (strongly disagree to strongly agree) to statement about the foundations of the Year 1 Checkpoints.

**Table 4.11:** Distribution of Responses across the 7-point Likert scale on the extent to which the February assessments were judged as effective by Prep teachers.

**Table 4.12:** Assessment choices of Year 1 Teachers in June Checkpoint.

**Table 4.13:** Distribution of responses (n & %) across a 7 point Likert scale when asked to comment on the usefulness of components of each Assessment in the Diverse Family June Checkpoint.
Table 4.15: Number and percentage of respondents who implemented each of the October Assessments.

Table 4.16: Distribution of responses across the 7 point Likert scale for Shadows are Everywhere, when asked to comment on the usefulness of each component.

Table 4.17: Distribution of responses across the 7 point Likert scale for Shopping when asked to comment on the usefulness of each component. Shading represents the most popular code selected.

Table 4.18: Distribution of responses across a 7 point Likert scale when Year 1 teachers were asked to comment on their level of agreement (strongly disagree to strongly agree) on the use of the October Assessments as a diagnostic tool, or more generally as an early years assessment process.

Table 4.19: Distribution of responses across a 7 point Likert scale when Facilitators were asked to comment on their level of agreement (strongly disagree to strongly agree) on the use of the October Assessments as a diagnostic tool, or more generally as an early years assessment process.

Table 5.1: Distribution of Participants who responded to Moderation Survey across Moderation Sites

Table 5.2: Distribution of Responses across the 7-point Likert scale on level of agreement with statements about the approach to moderation.
LIST OF FIGURES

**Figure 4.1:** Distribution of responses on the 7-point Likert Scale of all survey participants who reported having attended an on-line training session.

**Figure 4.2:** Distribution of responses of all survey participants who reported participating in an on-line discussion on the 7-point Likert Scale when asked to rate the usefulness of the forum.

**Figure 4.3:** Distribution of responses of all survey participants who reported contacting the QSA team for support on the 7-point Likert Scale when asked to rate the usefulness of the contact.

**Figure 4.4:** Distribution of responses of all survey participants who reported contacting the other EQ staff members for support on the 7-point Likert Scale when asked to rate the usefulness of the contact.

**Figure 4.5:** Distribution of responses across a 7 point Likert scale of the usefulness of the Implementation Plan – Reading within the Diverse Family Assessment represented as percentages.

**Figure 4.6:** Distribution of responses across a 7 point Likert scale of the usefulness of the Implementation Plan – Writing and Designing within the Diverse Family Assessment represented as percentages.

**Figure 5.1:** Distribution of responses in % on the 7-point Likert Scale

*The Moderation Day supported my work in assessment of Year 1 students.*

**Figure 5.2:** Distribution of responses on the 7-point Likert Scale

*The approach to moderation taken during the session was effective for this purpose.*

**Figure 5.3:** Distribution of responses on the 7-point Likert Scale

*The process of facilitation was effective.*

**Figure 5.4:** Distribution of responses on the 7-point Likert Scale

*The examples in the workbook helped to support my learning or the learning of other moderation group members about making consistent judgments.*

**Figure 5.5:** Distribution of responses on the 7-point Likert Scale

*The workbook activities helped to support my learning or the learning of other moderation group members about making consistent judgments.*

**Figure 5.6:** Distribution of responses on the 7-point Likert Scale

*The think, pair, share approach to arriving at a consistent judgment was effective.*
**Figure 5.7:** Distribution of responses on the 7-point Likert Scale

*The example analysis table and DAAR helped to support my learning or the learning of other moderation group members about acting on the evidence and future directions.*

**Figure 5.8:** Distribution of responses on the 7-point Likert Scale

*The activities related to the DAAR and acting on the evidence helped to support my learning or the learning of other moderation group members.*

**Figure 5.9:** Distribution of responses on the 7-point Likert Scale

*The moderation day was generally useful.*
1. Introduction

The Problem
Despite a focus on assessment, accountability and standards-based reform across the past decade in countries like Australia, and moves toward national tests, league tables, consistent curriculum content, standards and assessment regimes across Australian states and territories, the achievement of a balanced, rigorous debate about assessment has been somewhat elusive. Initiatives that have foregrounded assessment and accountability within narrow frames of testing have not worked to build the assessment capacity of teachers, nor it seems have they sustained improved outcomes for students (Luke & Woods, 2009). In such a context, systems’ approaches to school-based assessment that are rigorous, valid, and focused on: developing the professional assessment capacity of teachers; linking assessment with subsequent teaching and learning; and providing a more nuanced and fulsome version of student achievement and school outcomes than is possible as a result of standardised tests are crucial. This report evaluates an attempt to take such an approach.

Entry or early assessment of young people in the first years of schooling does not have the same history in Queensland as in some other systems. New Zealand, for example, has a tradition of assessing children using Clay’s Observation Survey (Clay, 1993) as they turn six and in their first year of schooling. The assessment tool is used as a means of identifying future intervention needs and planning learning and teaching. In Queensland, the introduction of the Year 2 Diagnostic Net (Department of Education, 1996) in the mid nineteen nineties, with the expectation that teachers in years 1, 2 and 3 would:

- Monitor students’ literacy and numeracy development;
- Validate judgments made using standard validation tasks in Year 2 and engage in a process of moderation as a way to identify students for additional intervention; and
- Report to parents on student literacy and numeracy achievement,

was a first move toward a systems’ approach to school-based assessment in the early years of schooling. The Year 2 Diagnostic Net was introduced as a result of the release of the 1994 Shaping the Future Report (Wiltshire, McMeniman, and Tolhurst, 1994) and the Literacy and Numeracy Strategy: 1994-98 (Queensland Department of Education, 1994). It has continued to be the cornerstone of early childhood assessment and reporting practice in Qld since its inception in 1995. There have been no substantial changes made to either the continua used as a framework for describing student achievement, or the validation tasks since its introduction. There have been some changes
made to expectations, requirements and some structures in the fifteen years of its use. These adjustments seem to have been made largely for pragmatic and administrative purposes (van Kraayenoord, Luke, Elkins and Land, 1999). This is despite questions being raised about the Net’s effectiveness and appropriateness as a result of several reviews.

The current early years’ context has seen the introduction of first state-wide, and more recently national, testing programs in literacy and numeracy in year 3 (also years 5, 7 and 9). There is already evidence that there are implications of the introduction of the National Assessment Program in Literacy and Numeracy (NAPLAN) for teachers’ assessment and pedagogic practice (see for example Comber, in press). *In such a context it is important that systems be involved in the design, resourcing and implementation of a systems-based approach to school-based assessment programs that promote consistent teacher judgments, effective assessment and pedagogical practice in the best interests of all students.*

The map of literacy and numeracy assessment programs across the primary year levels in Queensland continues to become more complex. Achieving balance, variety, and purpose continue to be key challenges. There are expectations of teachers to be involved in the implementation of the Early Years Record (Prep), the Year 2 Diagnostic Net (Years 1-3), the National Assessment Program in Literacy and Numeracy (Years 3, 5, 7, & 9), and the Queensland Comparable Assessment Tasks (Years 4, 6 & 9). Teachers in the early years of schooling are involved in the first three of these initiatives. What is important here is the impact of this assessment program on improved student outcomes.

So the questions for this evaluation are not about whether teachers can implement the Year 1 Checkpoints Assessments. We know that early years’ teachers in Queensland can, and have in the past, taken on the tasks of mandatory assessment programs, eventually embedding them into their routines and practices, and in many cases shifting those routines, practices and approaches as a result of the new assessment task. The questions that need to be asked are:

- How can a suite of assessments such as the Year 1 Checkpoints be used diagnostically to inform future teaching in educationally useful ways?
- What do the Year 1 Checkpoints assess?
- What effective teacher assessment practice will the Year 1 Checkpoints enable?
- How do we ensure that the Year 1 Checkpoints remain flexible, and *in development* tools for the fair and equitable judgment of performance and standards for all students?

This report is not anti assessment. We do argue for a balanced approach to assessment that includes multiple measures (National Council of Teachers of
The Purposes of Early Years Assessment and Moderation in a Standards Based Reform Context

Assessment is the purposeful and systematic collection of information about students’ achievements (QSA, 2009 p. 1), teaching and learning programs and teaching. It should serve a variety of purposes, including:

- Promoting learning;
- Improving outcomes; and
- Informing both teaching and learning within schooling contexts.

A systems-based approach to assessment should work in the best interests of all students. An effective assessment approach should provide information to a variety of stakeholders such as: students, teachers, parents and systems. However these “consumers of assessment data should be knowledgeable about the things the ... data can and cannot say about learning” (NCTE, available at http://www.ncte.org/positions/statements/assessmentframingst) and learners. However one tool or instrument should not be expected to ‘fit’ all of these purposes. What is called for is an approach across the early years of schooling that is consistent and coherent.

Generally large-scale assessment, and the practices of assessment in systems, assume that what is tested or assessed reflects the intrinsic abilities, capacities, and/or knowledge base of individuals. That is, there is an assumption that assessment is free of bias and somehow measures an absolute. This is contrary to a recognition that all assessment is the product of, and thus a reflection of the beliefs, understandings and knowledge of a particular community or culture.

Assessment is generally broadly construed under three categories - diagnostic, formative or summative. Much attention in the literature has focused on the difference between summative assessment – or assessment of learning – and formative assessment – assessment for learning – where students’ performance on tasks is used to inform planning, teaching and learning for individuals and groups of students. Diagnostic information can
and is taken from both summative and formative assessment activities, although it is most likely to be linked with formative assessment purposes. Any assessment instrument or type can, and often does involve more than one of these elements. Regular classroom based assessment can be classified as formative in that it provides the opportunity for students to receive feedback on their progress but this same work can also accumulate and count towards a summative statement of achievement. Assessment primarily aimed at diagnosis is intrinsically formative but it is possible to use the information collected via a summative assessment tool diagnostically (Harvey, 2009). As Harlen (2005) points out “the same information, gathered in the same way, would be called formative if it were used to help learning and teaching, or summative if it were not so utilised but only employed for (the purpose of) recording and reporting” (Harlen, 2005, p. 208).

Biggs (1998) contends that whether or not formative and summative assessments are seen as mutually exclusive depends entirely on how inclusive the model of assessment being used is. Harlen (2005) adds that the “...distinction between formative and summative purposes of assessment should be maintained, while assessment systems should be planned and implemented to enable evidence of students’ ongoing learning to be used for both purposes” (p. 207).

**Summative Assessment**

Summative assessment is often used to gauge a student’s level of achievement in relation to a particular program of study or unit of work. The term product orientated assessment is often associated with summative assessment. Harlen (2005) conducted two systematic reviews of research that supported the development of a definition of summative assessment by teachers as:

“The process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students’ learning, based on their professional judgment, and to report at a particular time on their students’ achievements.”

(Harlen, 2005, p. 213)

Summative assessment:

- Is usually conducted at the end of a unit of study to determine the level of understanding a student has achieved;
- Usually includes a mark or grade against a specified standard;
- Is frequently quantitative and often provides a comparison of a student’s achievement in relation to other students.

The summative uses of assessment can be either within the school itself or external to the school community. Within the school, results of summative assessments can be used to keep classroom grade books, and various other internal record keeping systems that are often used to report to parents and
students. In this manner, a range of school and classroom based summative assessment procedures can be considered as a part of an appropriate assessment system for young children. External use of summative assessments (most often, externally created tests or examinations) can include the monitoring of school performance and school accountability. When information from assessments is used to make important decisions about the performance of students, teachers and schools, the results acquire 'high stakes'.

'High stakes' summative assessments can impact not only on the learning experiences of the students but also on the nature of the assessments themselves – and this is of particular concern in the early years of schooling. When policy makers use assessment as a 'report card' by which to grade the performance of schools, there are consequent implications on curriculum, pedagogical practice and student's motivation for learning. High stakes assessment as accountability has been associated with teachers focusing on test content; teachers supporting students through repeated practice tests; teachers training students in the answers to specific types of questions; and teachers using transmission type pedagogical practices. It is likely that such practices can have indirect negative impacts on students' motivation for learning (Harlen, 2005).

The purpose of summative assessment is to grade children, and while this might be a component of what is necessary, if this purpose dominates the systems-based approach to early years assessment it is likely that the links between assessment and future teaching and learning will become more tenuous (Harlen, 2005). Variety of purpose is important.

**Formative Assessment**

Formative assessment is aimed at understanding and improving learning over time. It involves gathering and interpreting evidence during a course of study or unit of work. Formative assessment is designed to give students and teachers feedback on progress toward the development of knowledge, skills, understanding and attitudes rather than a point in time grade or mark. The outcomes of formative assessment can be used to determine the best and most appropriate teaching and planning decisions for the next step. Cowie and Bell (1999) define formative assessment as the bidirectional process between teacher and student to enhance, recognise and respond to the learning.

Formative assessment:

- Is used to determine a student’s knowledge, skills and understanding, including learning gaps as they progress through a unit of study;
- Is used to inform instruction and guide learning;
- Occurs during the course of a unit of study;
- Is embedded within instruction; and
• Is designed to give students feedback on their progress.

Formative assessment is considered integral to quality teaching and should involve students receiving immediate or timely and direct feedback. Feedback is to allow both students and teachers to gauge success, recognise areas of need and then plan future learning experiences based on this information. Feedback given is often descriptive and clearly points out areas of the student’s work that are successful as well as areas for possible improvement.

According to Boston (2002, p.5) some of the benefits of formative assessments for teachers include:

• “Being able to determine what standards students already know and to what degree;
• Deciding what minor modifications or major changes in instruction are required so that all students can succeed in upcoming instruction;
• The creation of appropriate lessons and activities for groups of learners or individual students; and
• The opportunity to inform students about their current progress in order to help them to set goals for improvement”.

The nature of formative assessment dictates that “a close relationship exists among curriculum content, instructional strategies and assessment” (Gullo, 2006, p. 140). Many aspects of early learning are best assessed as children engage in everyday, authentic experiences within their regular early childhood settings. Recent theoretical constructs that emphasise the socially constructed nature of learning and assessment remind us of the importance of the learning environment and the context in which learning and assessment take place. These constructs are particularly important when considering the quality of interactions between educators and young learners as assessments take place. Quality interactions within the everyday delivery of curriculum are increasingly recognised as central to all aspects of pedagogy, including assessment (Black and William, 1998).

Appropriate and effective use of formative assessment is not without challenges. Dunphy (2010) cites the three main issues of consideration here as:

• The professional knowledge of educators;
• Manageability in terms of time; as well as
• Educators’ sensitivity to the cultural and linguistic background of children.

Educators need an extensive understanding of early learning, as well as the theoretical constructs of formative assessment. Effective “assessment by classroom teachers requires a strengthening of teachers’ knowledge of what it is that, learned early, makes a difference to children’s continuing progression
in learning” (Department Education, Employment, Training and Youth Affairs, 1998, p. 19). Increasing an emphasis on learning in early childhood may be the best way to enact the field’s deepest commitments, and this approach can be implemented by providing programs and instructional experiences that improve all children’s chances for success now and in the future (Hatch, 2010).

**Diagnostic Assessment and Diagnostic Assessment Tools**

Diagnostic assessment is used to determine the nature of student learning in order to provide appropriate feedback and future teaching. One of the primary purposes of diagnostic assessment is to identify what students can and cannot do, which should lead to future planning and teaching, and the identification of students for intervention or other programs. Diagnostic assessment:

- Aims to improve learning and teaching;
- Can be systematically built into the curriculum;
- Identifies the next learning steps for students (and the next teaching steps for teachers);
- Must be linked to pre-determined learning objectives/milestones/expectations (the Framework);
- Could involve students in the identification of learning needs in the process; and
- Should involve the teacher in modifying the course or teaching strategies in light of the results

(Highland Council, 2010)

Forster (2009) makes a distinction between the following components of a **diagnostic assessment tool**:

- The **diagnostic framework** provides a structure for students’ learning to be recorded, monitored and reported.
- The **diagnostic instrument** is a tool or procedure that allows for the collection of evidence of students’ learning.

It is likely that the diagnostic tool also contains **adjunct teacher support** materials such as: statements defining the construct being assessed, or positions taken on student learning and effective teaching and assessment; suggested teaching strategies linked to performance on the instrument, annotated student work samples, and recording devices. The quality of these adjunct resources is implicated in the quality, effectiveness and utility of the tool overall.

Once a diagnostic tool is recruited for use within a systemic approach to assessment, or utilised for the purposes of student identification and
subsequent allocation of resources, or evaluation of teacher, school, program or system efficacy, there is an added dimension required. **Expectations** of teachers, schools and students must be clearly laid out in a coherent format, and this requires considerable reflection on the part of systems in relation to the consequent implications of this standard setting process. Additionally a quality diagnostic tool, utilised effectively within a systems based approach must necessarily clearly state the **delivery standards** (Luke, Weir, & Woods, 2008) attributable to the success of the process more broadly.

The question that need to be answered by systems then goes further than merely to ask about the Year 1 Checkpoints as assessment tools or even as diagnostic instruments. The diagnostic framework – The Year 1 Indicators – while not part of the brief for this Evaluation requires attention. Much more broadly than merely querying the quality of the tool, is the question of what the implications of embedding these materials within a systemic approach to early years assessment might be. It is in the interests of the system, schools, teachers, and students that our approaches to assessment are valid and equitable. That our approaches to assessment enable all students, regardless of their language, cultural, economic or racial background, the opportunity to perform at their potential, and to receive quality teaching and learning as a result of participating in the assessment seems logical. It is our belief however that this is not always the case, and as such within this report we argue for a measured, reflective approach to decisions about how the Year 1 Checkpoints will be used in the future, and the implications of these decisions for systems and schools.

In this way, the research reported here argues for a broad perspective and equitable, fair, and valid processes of early assessment in Queensland schools. We recommend a slow and thoughtful approach to the implementation of the Checkpoints regardless of the form that that implementation may take. We call for a focus on a systems-based approach to early years’ assessment that takes account of initiatives already in place and works toward a variety of useful purposes, means, and uses of the assessment suite. We call for lessons learnt from fifteen years of using the Year 2 Diagnostic Net to be carried forward, and issues identified over time, to not be attributed solely to the individual diagnostic tool in question. If we are to work in the educational interests of all students – including Indigenous students, low SES and those with particular issues related to language, learning, physical, cultural and cognitive capacities – we must ensure our approaches to early years assessment are effective, fair and accurate.
2. Considerations for Assessment

Introduction
Effective assessment practice is linked to quality planning and teaching and improved student learning (Black and William, 1998). In Queensland there is evidence that teacher capacity in assessment has been improved by systemic mandates for particular assessment regimes in the past. As reported by van Kraayenoord, Luke, Elkins and Land (1999), the Year 2 Diagnostic Net processes worked as a pseudo professional development activity for early years teachers in the mid to late nineties as an example. Definitions of assessment vary but the term is generally used to describe processes of collecting and interpreting information about learning and teaching. It should provide information on students’ progress, levels of achievement and performance standards. To be effective it must be utilised to inform future teaching and learning (New South Wales Department of Education and Training, 2007).

A balanced assessment approach may include different assessment for different purpose such as:

- Diagnostic assessments for identification of special needs and programs, assessment to support learning;
- Formative assessments to inform and modify teaching and learning;
- Summative assessments for high-stakes accountability purposes, program evaluation and monitoring trends

(Shepard, Kagan and Wurtz (1998)

No single assessment item can provide the breadth of assessment information required for all these purposes. Therefore, a balanced approach to assessment calls on a variety of assessment tools closely matched to the intent of the assessment purposes. Stiggins (2005) supports this as he indicates that the perfect assessment system would seek to provide a balance of assessment purposes and types. That balance would see a wide range of assessment for learning opportunities to assist students learning. Additionally, formative assessments would help teachers to view student progress toward standards and allow teachers to gauge which students need more assistance. And finally summative type assessments as well as the point in time accountability type assessment techniques could serve to verify the range of successes.

While summative assessment is often linked to a notion of assessment of learning, and formative assessment to assessment for learning, current
assessments understandings also consider *assessment as learning* as an important concept. In the sections that follow each of these concepts will be defined and considered in light of the unique nature of development in early childhood education.

**Assessment of Learning**

Assessment of learning is designed to measure student achievement and gauge what they have learned. As a measure of accountability, NAPLAN is such an assessment and aims to indicate how well schools and their students perform academically.

Assessment of learning:

- Is accompanied by a grade, score or mark;
- Evaluates the extent to which students have developed their knowledge, understanding and abilities in relation to set standards;
- Usually occurs at the end of a unit of work or stage of education; and
- Uses information gathered for reporting.

Assessment of learning used for accountability purposes is usually configured as some form of standardised testing. The logic of the testing as accountability movement can be linked to the publication of "A Nation at Risk" (National Commission on Excellence in Education, 1983). The report indicated ‘a rising tide of mediocrity’ in American schools. Similar trends can also be traced in the United Kingdom and Australia over the past decade.

The underlying assumptions of assessment of learning for accountability purposes have been criticised as being particularly problematic for early childhood contexts. Schafer (2002 as cited in Chan and Wong, 2010) indicates that there is usually a lack of cohesion between the curriculum, the learning activities, and the assessment items. Smith, Kuhs and Ryan (1993) report that the curriculum goals and instructional approaches inherent in standardised testing regimes are not usually consistent with young children’s developmental needs. She contends that the overemphasis on standardised testing can lead to instructional activities such as increased drill and practice and memorisation of isolated skills that are believed to be inappropriate in early childhood contexts.

The limitations and inappropriate nature of standardised testing with young children have been described by a number of educators (Gnezda, 1991 & Kamii, 1990 as cited in Smith et al, 1993). Standardised testing can yield limited and sometimes inaccurate information about learners including invalid measures of children’s abilities and potential. This information can then be used for the educational purpose of labelling, sorting and tracking young children into groups where the educational experiences provided are prescribed according to their assumed potential. The cultural fairness of these assessment tools has been called into question over decades, from early

Hatch (2010) argues that academic achievement does not equal learning and that children's learning, not test scores per se, should be the driving force behind teacher decisions about curriculum and assessment.

We are arguing here for appropriate assessment that is be geared toward the developmental, social and individual characteristics of young children to provide a more complete, accurate and comprehensive picture of a young child's learning. Assessment practices that are consistent with classroom instruction and curriculum goals are called for (Smith et al, 1993). Classrooms should be places where learning goals are fostered and assessment should work to let children know that learning itself is what really counts (Hatch, 2010).

Assessment for Learning

Assessment for learning occurs when teachers use inferences about student progress to inform their planning and teaching. Assessment for learning refers to the full range of assessments that teachers carry out on a regular basis to collect information about students' learning to inform decision-making about both teaching and the next step in a student's learning.

Assessment for learning:

- Provides feedback to students;
- Emphasises strengths, identifies student challenges and directs to the next steps for learning;
- Assists teachers to check on student understanding to inform future planning and teaching;
- Occurs as part of the learning process, from the outset through to possible summative assessment at the end of the study unit;
- Generally does not award grades or scores, but rather records achievement of milestones or describes quality against standards; and
- Can call on a variety of assessment items and types.

Recent perspectives on learning emphasises the social nature of learning and development in learning in the early years, and assessment of learning that is out of context can fail to recognize these established understandings of the social nature of learning (see for example Dunphy, 2010). Flewitt (2005, as cited in Hatch, 2010) indicates that young children co-construct meaning with adults using a range of communication techniques including facial expression and body movements. These behaviours often accompany talk and
supplement young children’s linguistic resources and abilities. The implications of this for assessment, are that teachers need to consider the multi-modal nature of children’s expression and attempts at meaning making. It is essential therefore, that early childhood assessment methods consider these issues as part of the unique nature of development in early childhood. The complexity of early learning dictates the employment of assessment methods that will allow for the development of suitably rich and accurate accounts of children’s learning.

Understandings of assessment for learning enable teachers to take into consideration the nature of learning in early childhood settings. Given all that is known about how children acquire knowledge and the best ways to teach, it is imperative that curriculum, instruction and assessment are considered inseparable components of the education process. Hatch (2010, p. 265) indicates that “assessment should be part and parcel of every learning activity, and opportunities for feedback and re-teaching should be built into all learning activities. Assessment need not be separated from curriculum and teaching, and it need not be sterile or threatening.”

Assessment used in this manner also allows educators to gauge not only what young children are able to do independently but also what they are capable of with assistance and what type of assistance is of most benefit. It enables the teacher to assess the impact of various levels and types of support on the child’s progress (Dunphy, 2010.) In addition, information gleaned from assessment for learning “directly feeds back into the teaching-learning process, letting the teacher know whether the children have learned the content at an independent level, whether they can demonstrate competence with the assistance of the teacher or others, and what kind of scaffolding works best to support students who cannot do the task without assistance” (Hatch, 2010, p. 265).

Assessment in the early years of schooling should be designed to meet the contextual needs of the school and students and be embedded in the local curricula (Chan and Wong, 2010). Assessment embedded in the practices of teaching and learning in the classroom can enable “an integrated collection of authentic data that truly reflected what a child really knows and what tasks the child can perform” (Chan and Wong, 2010, p. 9). Such assessment is more likely to promote “developmentally and culturally valid assessment of young children’s learning” (Chan and Wong, 2010, p. 9). Such assessment will also support the improvement of teaching and learning practices.

To summarise, assessment for learning provides for:

- The provision of effective feedback to students;
- The active involvement of students in their own learning;
- Opportunities to modify teaching and learning programs responsively; and
• Student self assessment, and an awareness of learning needs within the students themselves.

**Assessment as Learning**

When the principles of assessment for learning are fully engaged in classrooms, opportunities also arise for assessment as learning. Assessment as learning is that assessment that allows for students to reflect on the outcomes of assessment as a way to inform their own future learning goals and achievements.

Assessment as learning:

• Is ongoing throughout the learning process;

• Supports students in becoming aware of the goals of instruction and is transparent about the criteria for successful performance on any task;

• Involves the students in goal setting, monitoring progress and reflecting on results; and

• Implies student ownership and responsibility for learning progress.

The notion of assessment as learning is based on the premise that student achievement and academic self-concept are determined largely by the students’ perceptions of their own classroom successes. According to Stiggins (2000) it presents a philosophy that puts students at the centre of the classroom processes. Students become consumers of their own assessment information. They use evidence of their own progress to understand what comes next for them in the learning cycle.

According to Stiggins (2005) the process here is one of providing students with a clear vision of the learning target from the beginning of the learning process. Students see samples of high performance, along with examples that do not demonstrate understanding and this enables a more transparent understanding of the task, quality, and learning.

This notion of recognising, acknowledging and valuing the role that children themselves play in the assessment and learning process may, for some educators, present a new and challenging perspective on early childhood assessment. Dunphy (2010) believes that the challenges here relate both to professional preparation and structural issues of adult-child ratios.

To conclude, assessment of literacy and numeracy in the early years should be balanced across the instruments used, purposes, and styles deployed. It is legitimate to consider the necessary dimension of assessing students’ learning and its links to teaching and curriculum for accountability purposes. While certain types of assessment have been shown to be less than appropriate for young children in their first years of formal schooling, this does not take from the fact that young children and their teachers should be involved in assessment of, for and as learning. Assessment should be embedded into the pedagogic and curriculum practices of early years’ education. Data can be
collected on the literacy and numeracy achievement of our youngest school students. This data can be used across a variety of purposes, including making judgments about teacher practice and expectations, the quality of programs, the nature of student learning, and the identification of students’ needs. It would seem unlikely that one assessment tool – no matter how good – can achieve all of these aims. *In light of this, the purposes to which the Year 1 Checkpoints are deployed will have consequences for their effectiveness, quality of outcomes, and uptake.*

In what follows we examine the contexts and background of current early years’ assessment across systems in Australia.
3. Early Years Assessment - Policy and Practice in Australia

Introduction
In recent years, considerable attention has been given, in Australia and elsewhere, to the assessment of students early in their first years of schooling. This focus has often been as a result of the introduction of national agendas that provide frameworks for the improvement of students’ literacy and numeracy skills. As long ago as the late nineties, the National Literacy and Numeracy plan (Ministerial Council for Education, Early Childhood Development and Youth Affairs, 1997) prompted an investigation of early years assessments in 1998. The Curriculum Corporation, funded by the Commonwealth, carried out a collaborative project on current literacy and numeracy entry-level assessments (Curriculum Corporation, 1999). The report from this project identified a number of possible purposes for entry level or early years’ assessment:

- Obtaining a profile of the new entrant to be used as a basis for teaching, curriculum and classroom organization to meet individual needs;
- Identifying any children who may have difficulty and require early intervention;
- Developing a base line for measuring future progress in schooling;
- Assisting with resource allocation at the school and system level; and
- Provide system baseline information.

(Curriculum Corporation, 1999, p.1)

In addition Dever and Barta (2001) suggest two more possible purposes:

- Providing information for parents to support learning at home; and
- Providing reports to parents about student’s literacy and numeracy skills.

These purposes for assessment are in keeping with the foundations of the Year 1 Literacy and Numeracy Checkpoints Assessments, which also focus on:

- Providing direction for grouping students for effective teaching and learning; and
- Providing information to inform future planning and teaching of all students in the class.
Early Years Assessment

An extensive review of baseline or entry assessments was conducted by Sheila Wolfendale in the UK in the 1990’s (Wolfendale, 1993). Here she indicates that such early years assessments should always be in the best interests of the child, but also concludes that some form of on-entry or early to school assessment is educationally desirable. To follow on from this, in her editorial in the Journal of Research in Reading in 1999, she commented that:

If it is to be useful, baseline assessment must incorporate items which are developmentally valid, and give pointers to teachers on the next steps to take, including differentiation of the curriculum. Usefulness, however, is intimately related to the designated purpose, and one major issue for investigation will be the relative contribution of baseline assessment to identifying and planning children’s learning needs, in literacy and other domains, compared with its contribution to value-added analyses of the school...

(Wolfendale and Lindsay, 1999, p.6)

This perspective highlights several central issues for school entry or early years’ assessments. Firstly, the need for assessment to be developmentally appropriate, linked to students’ learning, and able to inform future teaching decisions. Assessments in the early years should also focus on the identification of individual students’ learning needs and classroom programs rather than reporting system-based data. Entry or early assessment should also be conducted in a meaningful context, allow for process to be assessed and not just product, be useful to teachers and part of everyday teaching and learning programs (Scholl Entry Assessment (SEA) Guide for Teachers, 1997, p.6).

Many early childhood professionals advocate the need for assessing young children using multiple measures over time and in the context of classroom activities. This enables teachers to continuously adjust instruction to meet individual needs in the day-to-day experience of classroom learning. Such ongoing, performance-based assessment provides reliable and accurate information that could not be assessed by any single, point-in-time standardised assessment instrument (Dever & Barta, 2001).

The National Association for the Education of Young Children (NAEYC) indicates that it is the nature of young children’s individual emotional, physical, social and cognitive growth that requires a unique approach to assessment. The tremendous individual variation in timing and patterns of growth can mean that individual students may have changed significantly by the time the results from any standardised or centralised assessment have been aggregated and reported.

Standardised assessments, by definition are administered identically to each child. As such, teachers are unable to adopt any flexibility in order to address any developmental, individual and cultural needs of children in their class.
addition the reporting of results to state authorities heightens the potential for high-stakes of the assessment.

As a result, early childhood education literature indicates that observational assessment of children’s progress; collection and annotation of work samples; and documentation of their development over a period of time are the most appropriate assessment practices. Observation of students as they complete a task provides a richer picture of their capabilities and is considerably less stressful for young children in particular. Assessment should be embedded in instruction over a period of time (Vukelich, 1997, as cited in Dever & Barta, 2001). Decisions that have a major impact on children should be based on multiple sources of information and must take into account the process of development over time.

Entry-level or early years’ assessments should not be exempt from these considerations and should reflect the unique nature of early childhood and the key concepts from early childhood research. Entry or early assessments that aim to inform instruction and identify strengths and weaknesses in a child’s learning need to be carefully designed, delivered and implemented in order to provide a valid, reliable account of what a child is capable.

Entry-level or early assessments that align with most appropriate early childhood practice should:

- Collect information over a period of time;
- Collect a range of information;
- Include observation, work samples and documentation of development and growth;
- Be embedded in classroom practice and instruction;
- Be used to develop teaching programs appropriate to individual and group needs; and
- Identify the skills and understandings of young children early in their school careers, as a way to inform instructional planning and provide information to systems and parents are significant goals.

The challenge is meeting these objectives using approaches that are relevant, valid, reliable and supportive of what students know.

**The Australian Context:**

The current education reforms aimed at closing the gap in achievement in literacy and numeracy, and improving outcomes for all students have been driven by assessment and curriculum reform measures. The development of an Australian curriculum is on-going, national testing regimes in literacy and numeracy have been put in place, publication of schools’ results and data in the name of transparency and school choice have all worked to apply
pressure to teachers and teaching to take certain directions and not others. In such a context the importance of an assessment technique based in sound early years’ assessment practice is important. Taking a systems-based approach to early years’ assessment means that a complete assessment program for the early years of schooling can wrap around national initiatives such as NAPLAN providing a more balance and fit for purpose.

In 2008, the then new Prime Minister announced an *Education Revolution* and subsequent policy has focused on literacy and numeracy, *closing the education gap* and increasing participation in new technologies. Part of recent policy to achieve these goals has included the signing of the National Partnerships Agreement (Council of Australian Governments, 2008) and a subsequent prioritising of:

- Effective and evidence-based teaching of literacy and numeracy;
- Strong school leadership and whole school engagement with literacy and numeracy; and
- Monitoring student and school literacy and numeracy performance to identify where support is needed.

(Council of Australian Governments, 2008, p. 3)

Currently, while this policy landscape has implications for early years’ assessment, states and territories still have the ability to, and continue to take the opportunity to, develop a range of assessment tasks that reflect the diversity of experiences and the social, cultural and economic differences that are inherent in each state or territory. As a result a variety of assessment tools and systems remain in place, and are under continued refinement or reform, in Australian schools and systems.

Currently, Australian states and territories use a range of terms to identify the first year of schooling, and there is a variety of minimum entry ages, and a number of different entry or early years’ assessment techniques and tools being used. Table 3.1 below outlines the name of the first year of school and the required starting age for each Australian State and Territory. Table 3.2 provided as a fold out insert to this report, provides a summary of the assessment items used in each system along with other associated information.
We now briefly detail the approach to early years’ assessment taken within each state and territory in the section that follows.

**Northern Territory**

The first year of schooling in the Northern Territory is referred to as transition. Commencing in 2011, transition students will be assessed using two tools, the Assessment of Student Competencies (ASC) screening assessment and The Prioritising Literacy and Numeracy Diagnostic Net for Transition to Year 9 (T-9 Diagnostic Net).

The ASC is a diagnostic tool used to measure those outcomes considered critical for success in successive, formal academic years of schooling (from year one onwards). The ASC is a competency based screening tool that requires a yes/no response from the classroom teacher on a range of indicators. Whilst some of the indicators are aligned with the Northern Territory’s Curriculum Framework, the competencies measured via the ASC only identify the *minimum* requirements considered critical for a child entering compulsory, formal schooling.

Whilst the ASC is mandatory for transition students, it is permissible to begin assessing students in their final term of preschool. Students who have demonstrated attainment of minimum standards by the end of preschool do not require retesting in their transition year. Students should be assessed using the ASC by Term 1 Week 5, and retested throughout the year as they attain individual competencies.

The ASC was developed locally by the Department of Education and Training in the Northern Territory. The approach is part of the Prioritising Literacy
and Numeracy: A strategy to improve literacy and numeracy outcomes 2010–2012 (NT Dept of Education and Training, 2010) initiatives, and was devised to comply with targets set as part of The Aboriginal and Torres Strait Islander Education Action Plan 2010–2014 (MCECDYA, 2010). It was designed specifically to take into account local demographic and socio-economic factors, such as the large Indigenous and remote student populations. A range of national and international assessments were used as reference during its development.

The ASC assesses four areas:

- Health and physical education including perceptual motor skills and fine and gross motor;
- Some Northern Territory Curriculum Framework EsseNTial Learnings;
- Early literacy aspects in listening, speaking, reading and writing;
- Early numeracy aspects in number, space, measurement and data;

and aims to:

- Provide baseline data on student learning;
- Inform and guide targeted teaching programs;
- Identify at risk students early in order to develop suitable and timely support programs to assist students;
- Report on student learning at a variety of levels (class, school and system);
- Provide parents and educators with a clear picture of student skills.

(Northern Territory Department of Education and Training, 2010)

Teachers are advised to take opportunities to assess students as part of the teaching and learning planned within daily classroom programs. As such the screening tool is to be used as an ongoing classroom assessment resource. Possible opportunities to assess students include observations of student behaviour or performance and collection of evidence of learning (work samples, photos etc). Teachers are required to enter the results of all assessments onto the ASC database. Data gathered and recorded using the ASC screening tool aims to inform teachers of the immediate learning needs of their students.

Teachers are required to develop early intervention (such as focused teaching episodes) for students where competencies are not yet demonstrated. The Teacher Handbook that accompanies the assessment materials outlines a range of teaching strategies for each competency. Students who, at the end of the transition year, have a significant number of ‘no’ responses recorded (in particular who do not meet two competencies deemed to be ‘critical’) are considered ‘at risk’. For these children, the classroom teacher is also
encouraged to involve a Special Education teacher and a range of support personnel in the design and development of an education adjustment plan for the student. The materials also recommend that a child’s family be included in discussions at this stage.

From 2011, the Northern Territory is implementing a new diagnostic net entitled the *Prioritising Literacy and Numeracy Diagnostic Net for Transition to Year 9* (Department of Education and Training, 2011). The use of this net is a mandatory assessment for students from Transition to Year 9. There are 3 Literacy continua (Reading, Writing and Oral Language Development) and the Numeracy continua (Numbers and how they work, Operating and Calculating, Shapes and Measurement, Time, Chance and Data, and Locations and Maps). Students who do not demonstrate competencies in the ASC will be unlikely to meet expectations in the Diagnostic Net. Like the ASC, the T-9 Net is a set of minimum standards considered necessary in order for students to successfully meet the demands of the successive year of schooling. It does not claim to encompass all literacy or numeracy requirements within the curriculum.

**Victoria**

The first year of schooling in Victoria is referred to as Preparatory or Prep. Victorian government schools are currently using the English Online Interview (Department of Education and Early Childhood Development, (Victoria) 2010) with a range of modules for different year levels, as a mandated assessment for students from the preparatory year through to year 2. The development of the English Online Interview was a combined project between the Department of Education and Early Childhood Development and Australian Council of Educational Research (ACER). A pilot of the system was trailed in 2008, and tests for the validity and reliability of the materials and process were conducted by ACER in 2009 (for Modules 2, 3 and 4 at the end of the preparatory year) and 2010 (module 1 at the start of preparatory year). In 2011 the English Online Interview assessment period will be from February 4 to March 4. During this time, all students in Prep, year 1 and year 2 will be assessed (Department of Education and Early Childhood Development, Victoria, Australia, 2010 (b), p.5)

Each module includes tasks and questions across a range of difficulty that match the expected range of student achievement for students beginning years Prep to year 2. The recommended modules for each year level are:

<table>
<thead>
<tr>
<th>Year Level</th>
<th>Recommended Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start of prep</td>
<td>Module 1</td>
</tr>
<tr>
<td>Start of year 1</td>
<td>Module 2</td>
</tr>
<tr>
<td>Start of year 2</td>
<td>Module 3</td>
</tr>
</tbody>
</table>
The English Online Interview is a ‘point in time’ assessment. It is conducted as an interview between the teacher and the student with responses being recorded by the teacher in an on-line system. The Interview assesses students against all dimensions of English Victorian Essential Learning Standards (VELS) (reading, writing, speaking and listening), and claims to provide comprehensive diagnostic information about student achievement in these areas. Groups of students with similar teaching needs can be identified based on their responses, and patterns and trends of learning within a class can be identified.

According to the Victorian Department of Education, the English Online Interview is aligned with both the Victorian Essential Learning Standards in English dimensions of reading, writing, speaking and listening and with NAPLAN.

In addition to the English Online Interview, Victoria also uses a Mathematics Online Interview. The interview was developed as a part of the Early Numeracy Research Project (1999-2001) and consists of a range of hands-on assessment tasks where students demonstrate mathematical understanding and preferred strategies for solving increasingly complex tasks. With a focus on mental computation the interview assesses in the areas of counting, place value, use of strategies for addition, subtraction, multiplication and division, time, length, mass, shape and visualization. (Department of Education and Early Childhood Development, Victoria, Australia, 2010)

**Western Australia**

The first year of schooling in Western Australia is referred to as Pre-primary. Western Australia is currently using the Online Interview as used in Victoria to assess the early literacy capabilities of students. In term 1 of 2010, Phase 1 was undertaken for Literacy Module 1. In this instance 50 schools delivered the online interview and 117 schools delivered the same interview in hard copy. Data and feedback collected from this study informed Phase 2 in 2010. Full cohort inclusion has begun in 2011.

In addition to the English Online Interview, Western Australia also trialled the Mathematics Online assessment tool, with items taken from the Victorian assessment tool of the same name. In term 2 of 2010 data was collected from 400 students to assist with the development of items for Western Australia’s Numeracy Modules. All pre-primary students will eventually be assessed using Module 1 in both literacy and numeracy. It is also expected that teachers of year 1 students will begin to use materials developed for Literacy Module 2, however this process is yet to be mandated.

As in Victoria, the assessments are in an interview format. Teachers use an on-line facility to access instructions and record student responses and processes. Some components of the assessment include ‘hands on’ tasks that require the use of books, stimulus sheets and counting and measuring items. The assessments target early essential skills in literacy and numeracy. The
The key purpose of the on-line assessments is the identification of literacy and numeracy skills critical to early educational success.

Literacy modules assess:

- Speaking and listening (aspects of oral language);
- Reading skills (phonological knowledge, phonemic awareness, concepts of print, reading, fluency, comprehension); and
- Writing.

Numeracy modules assess:

- Number (aspects of number and quantity awareness, counting, subitising, ordering, and one to one correspondence);
- Measurement (length and mass);
- Space (spatial awareness and shape).

According to the materials, the Online Interviews aim to:

- Identify students’ literacy and numeracy skills critical to early educational success;
- Provide valuable information to assist in the development of intentional and informed teaching for play-based programs, reflective of students’ needs;
- Provide information for early intervention of students identified as ‘at risk’ in critical aspects of literacy and numeracy;
- Provide teachers with group and class diagnostic reports for analysis, providing rich data to support class programs;

The reports generated for analysis are in the process of being linked to relevant content in the West Australian Syllabus, the First Steps program and the Australian Curriculum. (Department of Education, Western Australia, 2010 available at http://www.det.wa.edu.au/educationalmeasurement/detcms/education-measurement/articles/other-assessments/on-entry/on-entry-forms/frequently-asked-questions.en?oid=com.arsdigita.cms.contenttypes.FileStorageItem-id-10840385)

**Australian Capital Territory**

The first year of schooling in Australian Capital Territory (ACT) is referred to as Kindergarten. After what is detailed as a successful trial in 2000, ACT’s early years’ assessment program now uses Performance Indicators in Primary Schools (PIPS) to assess early reading, phonics and numeracy skills of students in the kindergarten year. PIPS baseline is a standardised assessment system designed to monitor students’ educational progress during their first year of schooling. It was established in 1991 and is offered by the Centre for
PIPS provides an assessment of attainment in the areas of reading, mathematics and phonological awareness. Assessments are mandated for all ACT kindergarten students in public schools and carried out in week 4 of term 1 and again in week 2 of term 4. By assessing students as they enter and again just before they leave, teachers are able to build a profile of individual progress for each child throughout their first year of schooling.

The class teacher delivers the assessment one-on-one with each student. Administration can be either computer based or completed as hard copy. As the child progresses through the assessment they see only those tasks required to identify their current skill level. PIPS as used in the ACT, aims to assess:

**Reading:**
- Handwriting;
- Vocabulary;
- Ideas about reading;
- Letter recognition;
- Word identification;
- Sentence reading;
- Construction of sentences.

**Phonics:**
- Repeating words – understanding of sounds and phonological awareness;
- Rhyming words – choosing of picture rhymes.

**Numeracy:**
- Ideas about maths;
- Counting;
- Sums A – solving simple number problems using small numbers (under 10);
- Numbers;
- Shapes;
- Sums;
- Maths.

The relevant materials claim that Pips is used in ACT aims to:
• Identify (as early as possible) students who may need extra support or extension in the areas of literacy and numeracy;

• Use data collected to assist teachers in planning appropriate learning experiences for individuals and class groups;

• Give teachers insights into how students determine answers to questions through the one-to-one administration of the assessment and the type and range of questions asked;

• Provide teachers with information on how well students have progressed during the year relative to their starting point;

• Predict future performance and assist teachers to identify students who might benefit from intervention programs;

• Have schools use data to monitor and revise teaching programs.

It is conducted by the classroom teacher in a classroom environment. It claims to be an integral component of the teaching and learning cycle and teachers are encouraged to use its results to inform teaching practices relevant to the needs of the individual child. Individual student reports are generated and provided to parents in term 2 and 4. The reports provide a numerical score for reading and numeracy, which then places the student within a 'band' for that test score. Scores are standardised and parents are provided with a table that highlights the percentage of students in each of the bands. The Centre for Evaluation and Monitoring at the University of Durham describes the assessments as 'curriculum-free' predictors of academic attainment. As such the PIPS system does not align directly with ACT curriculum or standards nor will it with the Australian Curriculum.

Tasmania

The first mandatory assessment of Tasmanian students occurs in the year prior to the first year of schooling. This year is referred to as Kindergarten. During the kindergarten year teachers monitor and observe skills in a range of areas (literacy and numeracy, social skills, fine and gross motor skills) using the mandated Kindergarten Development Check (KDC). Kindergarten students are assessed against 21 ‘critical markers’ in these areas during term 1. Critical markers are aligned to the Tasmanian Curriculum Framework and “are an early check for appropriate development”. Students who do not meet certain markers in term one are re-assessed in term 3. Information gathered during the KDC is used during parent-teacher discussions and reported in an end of year summary report (Available at http://www.education.tas.gov.au/earlylearning/earlyyears).

Much of the information collected by the KDC does not actually relate directly to literacy and numeracy skills, so in 2000 after a review of available tools, the Office for Educational Review suggested the use of the Performance Indicators in Primary Schools (PIPS) as an appropriate diagnostic tool for the
early years of schooling. The PIPS has been implemented with the full cohort of students since 2002.

The first year of schooling in Tasmania is referred to as Preparatory or Prep. During this year all students are now assessed using Performance Indicators in Primary Schools (PIPS). PIPS assesses early reading, phonics and numeracy skills of students in the prep year. PIPS baseline is a standardised assessment system designed to monitor students’ educational progress during their first year of schooling.

As used in the Tasmanian system, PIPS aims to assist in the identification of students who are not achieving expected standards in literacy and numeracy. PIPS consists of two separate assessments conducted in the first four weeks of term 1 and again in the first few weeks of term 3. Parents of students who score less than 40 in the first assessment are notified and teachers are tasked to develop an appropriate intervention program for these students. Children with scores below 40 in the second assessment continue to be provided with support in the following year. All parents of prep students are provided with information about their child’s performance on PIPS and have the opportunity to discuss results with class teachers.

PIPS as used in Tasmanian schools aims to assess:

Reading:
- Handwriting;
- Vocabulary;
- Ideas about reading;
- Letter recognition;
- Word identification;
- Sentence reading;
- Construction of sentences.

Phonics:
- Repeating words – understanding of sounds and phonological awareness;
- Rhyming words – choosing of picture rhymes.

Numeracy:
- Ideas about maths;
- Counting;
- Sums A – solving simple number problems using small numbers (under 10);
- Numbers;
• Shapes;
• Sums B - students may use pen and paper to answer e.g. “What is 2 more than 6?”;
• Maths – more complex maths such as finding the picture to match a simple fraction.

PIPS as used in Tasmanian schools aims to:
• “Inform learning and teaching programs and monitor individual student progress;
• Determine the starting point of the class;
• Identify individual strengths and weaknesses;
• Target specific students for individual programs;
• Group students for some aspects of their teaching.”


According to the system, data collected from PIPS assessments assists schools to monitor the impact of programs and make decisions about resourcing. The Department of Education, Tasmania uses PIPS data to provide entry level literacy and numeracy performance statistics for all students (including the data collected for the Indigenous Education Strategic Initiatives Program reporting) and to determine individual progress against Tasmanian norms. The Department also uses the information to identify the needs of specific sub groups and to meet the state and national goals for reporting requirements. (Tasmanian Department of Education Office for Educational Review, 2002, p.6)

South Australia
The first formal year of schooling in South Australia is referred to as Reception. South Australian teachers from Reception to year 3 are currently using the School Entry Assessment (SEA) as a process to collect information about a child’s progress in literacy and numeracy. The SEA claims to be an assessment instrument designed to support children to “demonstrate what they understand and what they can do in three key learning areas: early literacy, early numeracy and oral language”. First developed in New Zealand, where it became available for school use in 1997, it has been adapted for the South Australian environment by the Curriculum Corporation.

The South Australian Education Department introduced the School Entry Assessment (SEA) Policy in 2001 in an attempt to help teachers assess young learners and plan relevant learning events from Reception to Year 3. The implementation of the SEA policy included the development and publication
of the Learner Record booklet and accompanying Information for Educators handbook. The Learner Record booklet was designed to fulfil two key functions. Firstly it was to act as a tool for teaching and secondly it was to be used as a platform for communication between educational and social contexts. It incorporates five stages of learning that indicate a developmental continuum that supposedly helps to describe children’s ongoing learning. The stages of awareness, exploration, inquiry, utilisation and application are designed to allow educators to map literacy and numeracy growth as children move through the phases of learning. Entries are made in the Learner Record books in the last few weeks of a child’s first, fifth, ninth and thirteenth terms at school. During the fourth term of year 2 students are expected to have moved through all stages of the continuum and are then assigned a South Australian Curriculum Standards and Accountability (SACSA) standard. It is unclear what information is collected centrally through this process, or indeed which parts of the process are mandatory.

The SEA has three separate components: a reading task based on Concepts About Print (originally developed by Marie Clay); a numeracy assessment using a supermarket game ‘Checkout’ and an oral language task assessed via a storytelling activity called ‘Tell Me’. Some of the components that SEA aims to assess include:

- Concepts about print such as directionality, how words are made up and the correspondence between written and spoken words;
- Numeracy concepts including numeral recognition, forming sets, pattern recognition, number sequence knowledge;
- Communication skills including the early development of oral language skills.

According to Lees (2007) The South Australian Education Department state five aims in the School Entry Assessment (SEA) policy:

- The collection of information about the competencies and learning dispositions of beginning students within the first ten weeks of schooling (the actual school entry assessment part of SEA is mandated);
- The provision of a platform for home-school communication and communication between pre-schools, schools and other institutions;
- The provision of an ongoing record of each child’s development and progress towards Standard 1 in the SACSA Framework, with a focus on timely intervention to meet diverse needs;
- The provision of state wide data for research into better education practices and policies; and
- Informing teachers’ beliefs, planning and practices through ongoing training and supports.
The Department also outlines that recording the assessments of each child enables:

- “The development of classroom programs that include all strands of literacy and numeracy;
- The identification of children requiring additional support;
- The planning of appropriate programs for these children and an ongoing record of student development;
- The development of student reports”.

(Available at http://www.decs.sa.gov.au/ no page)

**New South Wales**

The first year of formal schooling in New South Wales is referred to as Kindergarten. In 2008, 438 schools in New South Wales commenced using the Best Start Kindergarten Assessment. It was developed locally by the K-4 Initiatives unit in conjunction with the literacy and numeracy units (for the State of New South Wales through the Department of Education and Training). The literacy component was developed in conjunction with ACER. The University of Newcastle reported on the first trial in 2007 and ACER trialled the assessments for validity and reliability during the trials in 2008 and 2009. In 2009 a further 600 schools implemented the program and all 1700 primary schools in NSW should have participated in 2010. From 2010 all primary schools will implement the new program which involves the administration of standard assessment tasks designed to identify the literacy and numeracy knowledge, skills and understandings that each child brings to school as they enter Kindergarten.

The Best Start Kindergarten Assessment takes place during term 1 of a child’s kindergarten school year. Assessment tasks are to be completed by the end of week 5 in term 1. Student data is then entered into software by the end of week 7 in term 1 and parent feedback is conducted in week 8 and term 1.

The Best Start Kindergarten Assessment aims to assess:

Seven aspects of literacy:

- Reading texts;
- Comprehension;
- Concepts about print;
- Phonics;
- Phonemic awareness;
- Aspects of speaking; and
- Aspects of writing.
And three aspects of Numeracy:

- Counting sequences;
  - numeral identification;
  - number word sequences;
- Counting as a problem solving process; and
- Pattern and the repeated unit.

The Best Start Kindergarten Assessment aims to provide Kindergarten teachers with information to build on students’ current understandings and develop quality teaching and learning experiences for the future by ascertaining the achievement levels of students.

**The Queensland Approach**

The first year of compulsory schooling in Queensland is year 1. Prior to year one, most students attend a non-compulsory year of education called Preparatory or Prep. Children’s learning is recorded on the Early Learning Record during the Prep year, and this information is passed to parents and year 1 teachers. As part of the Year 2 Diagnostic Net (Net) (Department of Education, 1996) procedures, children’s learning and development are monitored throughout the first three years of primary school (years 1, 2 and 3) using individual student profiles and developmental continua. The continua list explicit descriptors of behaviour that are claimed to support teachers in identifying how students are constructing and communicating meaning in reading, writing and numeracy. The descriptors, called indicators, are grouped into clusters called a ‘phase’. The clustering of indicators into phases is purported to allow teachers to map children’s overall progress across a number of phases. Teachers ‘map’ students onto the continua using a combination of observation, interview and collection of work samples.

In conjunction with the mapping of students, the Year 2 Diagnostic Net requires teachers to validate teacher judgments about students’ performance in May of year 2. These validation assessment procedures aim at confirming teacher judgments from the continua in reading, writing and numeracy and identifies year 2 students who require additional support. A sub-set of key indicators within each focus area (reading, writing and number) is used to determine which children will be involved in the validation process. Only those children not demonstrating selected key indicators are assessed in the point in time validation tasks. This is one of the key changes made to the Net after calls from teachers about excessive workload. Validation occurs in schools in term 2 of year 2 every year and is followed by a moderation process for teachers and Key Teachers. Students identified as requiring additional assistance are consequently supported by a range of programs within the school. This is resourced through intervention funding targeted at
specific individuals on the basis of their reported performance within the Year 2 Diagnostic Net. Teachers also use the Year 2 Diagnostic Net materials to complete standard written reports for parents of students in years 1, 2 and 3.

The developmental continua used within these processes are based on the Education Department of Western Australia’s *First Steps Developmental Continua* (1994a, b). The Year 2 Diagnostic Net was trialled in all Queensland state schools in 1995 and remains the foundation of system-based literacy and numeracy assessment in the early years. “The Year 2 Diagnostic Net aims to:

- Monitor and report on aspects of children’s literacy and numeracy development during the first three years of schooling;
- Identify those children who are experiencing difficulty in literacy and numeracy;
- Provide diagnostic information to the school authorities who support children experiencing difficulty;
- Validate teacher’s observations of year 2 children using specifically designed assessment tasks;
- Provide a report to parents about their child’s learning and development in aspects of literacy and numeracy”.

(Department of Education (Qld), 1996, p.1)

The Year 2 Diagnostic net aims to assess:

- Reading and how students:
  - Make meaning at text level;
  - Make meaning using content;
  - Make meaning at word level; and
  - Demonstrate attitudes towards reading.

- Writing and how students demonstrate:
  - Content, organisation and contextual understanding;
  - Concepts and conventions;
  - Strategies;
  - Attitude; and
  - Word usage, editing and language conventions.

- Number and students understanding of:
  - Counting and patterning;
  - Number concepts and numeration;
The process remains mandatory for all state schools.

**Previous evaluations of the Year 2 Diagnostic Net materials and processes**

The literacy and numeracy continua of the Year 2 Diagnostic Net originated in the Western Australian First Steps materials (1997). These materials have been evaluated several times since their first preparation, and have been redeveloped in answer to some of the critiques levelled over time (see Deschamp, 1994a, 1994b for example). It is true that several of the critiques of the materials in their original context are less relevant to the use of the continua within Qld's Year 2 Diagnostic Net. For example the criticism of the materials' reliance on teachers' decision making without a process for supporting consistent judgments has less weight within the Qld context, where year 2 teachers and their Key Teachers have been expected to moderate at school and cluster levels. However, critiques levelled at the lack of analysis of student achievement outcomes through the First Steps approach, particularly for Indigenous students or students from other ‘at-risk’ cohorts (these were the students for whom the First Steps program was first developed) should be taken into account in any investigation of the Year 2 Diagnostic Net while ever the materials form the basis of the Net process. During the latter part of the nineties the Northern Territory also utilised the First Steps materials as the foundation of their approach to improving outcomes for Indigenous students in literacy. Part of this uptake involved research and evaluation of the materials and approach (Northern Territory Department of Education, 1999). However, the evaluations aimed at documenting the impact and implications of the program as a whole, and did not evaluate the Developmental Continua or their effectiveness or validity.

The Qld Year 2 Diagnostic Net has also been reviewed on several occasions and as it draws close to sixteen years of service within Queensland schools calls for its renewal or replacement have continued. The Net process has been adapted throughout its service years¹ usually in answer to calls from teachers and unions, but there has been no substantial reform of the materials or the Net processes during the years of its implementation.

---

¹ As an example in 1995, the first year of the Net’s implementation, all students in Year 2 were validated as part of the Year 2 Diagnostic Net processes, however after calls from teachers and unions about workload, in 1996 a decision was made to limit the validation procedures to those who were expected to be identified as requiring extra assistance in the form of intervention. A sub-set of key indicators within each focus area (reading, writing and number) is used to determine which children will be involved in the validation process.
Relevant for this report are the past calls for systematic adaptations to the materials and processes which have for all intents and purposes gone unanswered. In 1997, the first of two reviews of the Year 2 Diagnostic Net processes was submitted but not released (Luke, Land, Elkins & van Kraayenoord, 1997). A published paper from the authors of the 1997 review in 1999 (van Kraayenoord, Luke, Elkins and Land, 1999) documented concerns with the basic assumptions at the foundation of the materials and process in relation to the views of: learning and literacy learning underpinning the approach; quality assessment; and categories of at risk student cohorts. The authors also expressed concerns about the relationship between the Year 2 Diagnostic Net materials and the then state based syllabus and curriculum documents in Mathematics and English and the potential for assessment driven narrowing of the curriculum, they claimed there were issues of validity and raised concerns about the moderation processes. There were also concerns about the how the data collected might be put to use in the political climate of the time.

Calling into question the model of literacy at the foundation of the Net materials and its currency for the nineties, the research team (Luke et al, 1997) claimed that there was a narrow focus which ignored important dimensions of literacy including phonics and critical literacy. There was a mix in the perspectives taken within the indicators of literate and numerate behaviour across the Net phases, resulting in what the researchers considered an ‘incommensurable’ conflict of models of literacy (and numeracy) development and learning. While the net was recognised as providing opportunities for students to demonstrate and teachers to collect, record and analyse evidence of literacy and numeracy learning within context, the individualised basis of the assessment was critiqued as ignoring the social practices generally recognised as important in literacy learning, achievement and assessment. Additionally the authors highlighted the ‘risk’ of being identified for certain populations of students. This was the first indication that the materials themselves might be implicated in the perpetuation of disadvantage and deficit understandings of students from diverse backgrounds. The research team called for a re-examination of the cultural content of the materials generally, and a re-analysis of the utilisation of monolingual single pathway taxonomies of student behaviour or learning – the continua - and their place in the production of inequitable outcomes.

After this first review (Luke, Land, Elkins and van Kraayenoord, 1997) was not released, a second review was commissioned (Stewart-Dore, Bartlett, Hallinan, Moreton, Robert and Woodrow, 1999). While this second team noted that the strict design limitations placed on them limited their ability to provide a close and specific analysis, they did raise concerns particularly about the linguistic differences of students in Qld schools, and the monolingual focus on Standard Australian English of the materials.
Six years after the first review was delivered, and as part of an investigation into cultural fair assessment within state-based assessment and accountability moves, a third round of calls for the renewal of the Year 2 Diagnostic Net materials and processes was made. Based on review of evaluations of the First Steps materials, and the Year 2 Diagnostic Net materials and processes, in Western Australia, Northern Territory, and Queensland a team from the University of Queensland (Luke, Woods, Land, Bahr & McFarland, 2003) called for further consideration of issues related to the mono cultural and mono lingual nature of the developmental continua and the Year 2 Diagnostic Net materials and processes. As part of this report there was a specific call for dedicated training for teachers in the use of the materials and processes in relation to diverse student groups², and for a review of the appropriateness of single pathway continua as a framework for making judgments about a diverse range of students. The calls from three reviews for the re-analysis of specific issues related to, or for changes to the Year 2 Diagnostic Materials and processes remained unanswered to date, and well over a decade since these calls began.

While this report is not an analysis of the effectiveness of the Year 2 Diagnostic Net, there are considerations from these past reviews of the Net for the investigation being reported here. *Issues related to cultural fair assessment, validity, and the importance of ensuring the ‘framework’ used to report and monitor students work is fair and inclusive of the possible pathways of a diverse range of students and requisite expectations need to be considered in the development of any early years diagnostic assessment for Queensland schools.*

**Conclusion**

All Australian States and Territories are currently in the process of trialling, moderating or implementing early years’ assessment tools or programs. In many cases these processes are new or are in the process of being renewed. The systems’ approaches are variable in scope, expectations, standards, and quality. Some use validated commercial style assessments, while other systems have developed assessments locally. The trade offs in validity and reliability are evident in these decisions. Queensland is still officially using a diagnostic tool first implemented more than fifteen years ago. The continua used as a framework to record students’ achievement was designed earlier again. The need to renew the approach to assessment in early years’ classrooms is now urgent. However any renewal of these assessment items must necessarily take account of the research and expertise that has built up as a result of Queensland’s approach to early years’ diagnostic assessment in

---

² The call was specifically made in relation to Indigenous students, as a result of the reports’ focus on culture fair assessment for Indigenous students.
the past two decades. *Any renewal efforts must consider much more than the replacement of one instrument with another.*
4. The Year 1 Checkpoints

Introduction

The draft Year 1 Literacy and Numeracy Checkpoints Assessments were in open and supported trial during Semester 2, 2010. The Checkpoints were designed as a diagnostic tool in 2009, as a way to incorporate the use of the Year 1 Literacy and Numeracy Indicators as formative assessment in Year 1 in Queensland Schools. The Supported Trial was externally evaluated as part of this trial. The Evaluation is the topic of this report.

In their current form the Year 1 Checkpoints provide assessment resources for teachers to use in February, June and October. They aim to support teachers in monitoring children’s progress and making judgments about their achievement of the targeted P-3 Literacy and Numeracy Indicators by the end of Year 1 (Queensland Studies Authority, 2010 p. 1).

These assessment resources have attempted to focus on the use of effective early years’ assessment practices where evidence is gathered from hands-on teaching and learning experiences, rather than more formal assessment methods. They are based in a model of assessment for learning, and aim to support teachers in the “ongoing process of determining future learning directions” (Queensland Studies Authority, 2010 p. 1) for all students. As such they work to encourage teachers to do more with assessment than identify students who are not achieving benchmarks or minimum standards, which has arguably been the focus of other early years’ assessment initiatives in Queensland. Instead they focus teachers on interpreting and analysing evidence to make informed judgments about the achievement of all students, as a way to support subsequent planning, teaching and learning.

The design of this section of the Evaluation was organised to combine a variety of data sources, through a variety of data collection techniques with the aim of gaining insight into as much detail of the Year 1 Checkpoints Assessments as possible. The data sources aimed to support the review of the materials with:

- observation of the Supported Trial decision making processes;

- quantitative data collected through four surveys:
  - the moderation survey;
  - the Year 1 teachers’ survey;
  - the Facilitators’ survey;
  - the Prep teachers’ survey; and

- qualitative data collected as part of surveys, interviews, written feedback provided by participants, and observations of Supported Trial events.
The purposes of the analysis that follows were to:

1. Evaluate the materials and report on the views of school-based staff involved in the trial on the process, materials, and assessment practices; and

2. Feed Forward considerations for the future use of the materials.

We look first at the Trial process, raising issues related to the resources and supports provided and the implication of time issues related to the truncated nature of the Supported Trial and the use of the assessment items. We then move to discuss the Year 1 Checkpoints Assessments materials more generally.

**Year 1 Checkpoints Assessments Trial Processes**

The Year 1 Checkpoints Assessment Supported Trial (hereafter the Supported Trial) has been conducted by the Queensland Studies Authority during semester 2, 2010. The Supported Trial selected 1003 Education Queensland schools from schools that had voluntarily registered their interest in taking part in the trial of the Year 1 Checkpoints materials. At least some of these schools had been involved in a smaller open trial in 2009. The sample schools represented schools across a variety of Education Queensland regions and included schools with:

- A high Indigenous student population;
- Rural and remote school locations;
- Single and multi-age early phase classes; and
- A high proportion of students from low SES backgrounds.

There were teachers of both multi-age and straight Year 1 and Prep classes teaching at the Supported Trial schools. Table 4.1 below provides a synopsis of the school categories within the trial.

---

3 One school subsequently removed themselves from the Supported Trial leaving ninety-nine schools as the sample.

52
Table 4.1: Number of schools in the Supported Trial according to relevant school categories.

<table>
<thead>
<tr>
<th>SES Status</th>
<th>Low SES</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9 Schools</td>
<td>90 schools</td>
</tr>
<tr>
<td>% Indigenous students</td>
<td>% Indigenous Students &gt; 10%</td>
<td>% Indigenous Students &lt; 10%</td>
</tr>
<tr>
<td></td>
<td>26 Schools</td>
<td>73 schools</td>
</tr>
<tr>
<td>Size of school</td>
<td>Small</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>18 schools</td>
<td>58 schools</td>
</tr>
<tr>
<td>Location</td>
<td>Metropolitan</td>
<td>Provincial</td>
</tr>
<tr>
<td></td>
<td>38 schools</td>
<td>13 schools</td>
</tr>
<tr>
<td>Region</td>
<td>South Qld</td>
<td>Central Qld</td>
</tr>
<tr>
<td></td>
<td>39 Schools</td>
<td>34 schools</td>
</tr>
</tbody>
</table>

Exact counts of Year 1 and Prep teachers involved in the trial are more difficult to ascertain. While schools originally nominated numbers of teachers who would participate at their school, communication from schools to QSA staff, accounts from teachers at moderation sessions, interview data from teachers, along with communication between schools and the Evaluators suggest that not all nominated teachers continued in the trial. This might be even more likely to be the case for Prep teachers, many of whom were not considered in the original nomination process, because the Supported Trial was named and considered to be a Year 1 teachers’ trial process. Consequently for many Prep teachers the information about requirements and expectations of being involved was delivered late in the year. What is clear is that while all schools bar one continued in the trial, not all teachers who were nominated fulfilled the requirements of being involved.

**Adjunct supports provided to Supported Trial Participants**

Teachers and facilitators involved in the Supported Trial received a variety of support mechanisms from the QSA and EQ. Along with copies of the materials, regular updates about the Supported Trial and the materials, on-line and face-to-face workshops at key junctures in the Supported Trial process, access to additional readings, support materials and financial support for a moderation day session for clusters of Supported Trial School staff were provided to those involved in the supported trial.

Data of the teachers’/facilitators’ responses to these support resources was collected through the Year 1 teachers’, Prep teachers’, and Facilitators’ surveys, interviews and written comments of participants. The data, which
provides illumination into the opinions of participants about the on-line training sessions; ongoing on-line discussion board, and phone/email support offered to participants on an ‘as needed’ basis are reported below.

**On-line Training Sessions**

Participants were asked to provide information about if they had participated in the on-line training sessions throughout the trial period, and those who did report being involved were asked to rate the sessions on a 7 point Likert scale. Seventy percent \( (n = 43) \) of Year 1 teacher participants, sixty percent \( (n=13) \) of Prep teacher participants, and ninety one percent \( (n=20) \) of Facilitators who responded to the surveys reported participating in on-line training sessions. Of all these participants, 16\% \( (n = 12) \) found the on-line training sessions extremely useful and 15\% \( (n = 11) \) identified with the middle of the scale. Only five percent \( (n = 4) \) reported finding the on-line training sessions not at all useful. Fifty-five percent of respondents rated the sessions above mid-point, with the largest group of participants \( (21.1\%) \) selecting a scale point 5. It is clear that there was a full range of responses to this technology format. This result must be seen in light of the fact that technical difficulties interrupted one of the three initial on-line training sessions. Figure 4.1 below represents the percentage of responses across a 7 point Likert scale when respondents who had participated in the sessions were asked how useful they had found those sessions. Table 4.2 provides percentages of respondents who found the training extremely useful, not useful at all and those who responded at the midpoint of the 7 point Likert scale.

![Figure 4.1: Distribution of responses on the 7-point Likert Scale of all survey participants who reported having attended an on-line training session.](image-url)
Table 4.2 – Numbers and percentages of participants who participated in the On-line Training Sessions and reported finding the sessions extremely useful, middle of the scale and not at all useful.

<table>
<thead>
<tr>
<th></th>
<th>Year 1 teachers</th>
<th>Prep teachers</th>
<th>Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td>% participated in on-line training</td>
<td>43 57.3</td>
<td>13 46.4</td>
<td>20 69</td>
</tr>
<tr>
<td>% found it extremely useful</td>
<td>3 7</td>
<td>5 38.5</td>
<td>4 20</td>
</tr>
<tr>
<td>% found it middle of scale</td>
<td>8 18.6</td>
<td>1 7.7</td>
<td>2 10</td>
</tr>
<tr>
<td>% found it not at all useful</td>
<td>3 7</td>
<td>1 7.7</td>
<td>-</td>
</tr>
</tbody>
</table>

The qualitative responses left by participants support the finding that while the percentages in the table above might be considered as evidence that there were differences in the respondents’ opinions of the on-line training sessions according to their teacher type, that this is not likely to have been the case. Qualitative comments across the three cohorts of Year 1 teachers, Prep teachers, and Facilitators of the process fit similar categories. Technological difficulties are discussed, however the opportunities that this technology afforded, particularly for those in remote areas is also highlighted. The technological platform used was both praised by some and discussed by others as unfamiliar and thus difficult and time consuming to use. Other relevant comments related to concerns with peer's ill preparedness for the sessions (*lots of the questions asked were answered in the materials anyway*). Comments ranged from those who felt the sessions clarified many issues, to those who commented that there was little new material covered. This suggests that the needs of participating teachers differed. While the location of teachers and the implications of this for receiving materials via postal services, or virtually also received comment – this must be considered within the context of the very short timelines of the trial.
CONSIDERATIONS:
Implementation plans must take account of time restraints, teacher location and time lag for delivery, technological difficulties, and environmental issues that might impact on access to materials and resources. There are time and resourcing issues consequent for teachers and schools when technology does not deliver. *Are there means to provide opportunities for alternatives in addition to ELUMINATE sessions?* Real-time virtual communication, while affording many benefits also have disadvantages, so provision of other forms – vodcasts with Q&A posted later for example – may need to be considered to supplement the advanced formats.

Consideration of teacher expertise level in the use of the Year 1 Checkpoint materials – particularly for those who have already participated and others who have not - may be required in the first few years of implementation and perhaps beyond.

**Ongoing On-line Discussion Forum**
Participants were asked to provide information about if they had participated in the ongoing on-line discussion forums throughout the trial period, and those who did report being involved were asked to rate the sessions on a 7 point Likert scale. Less survey respondents reported utilising this support than the on-line training sessions. Forty percent (*n* = 30) of the Year 1 Teacher sample, eighteen percent (*n*=5) of Prep teacher sample, and thirty-eight percent (*n*=11) of the Facilitator sample reported participating in the ongoing on-line discussion forums. Of all of these participants, 13% (*n* = 6) found this extremely useful, 4% (*n* = 2) found this not at all useful and 24% (*n* = 11) identified with the middle of the scale. Fifty-four percent of the participants selected above the mid-point, so again there is a full range of responses to this technology, and there are definitely teachers who found the forums useful. Figure 4.2 represents these results in diagrammatic form while Table 4.3 provides percentages of respondents who found the training extremely useful, not useful at all and who responded at the midpoint of the 7 point Likert scale.
Figure 4.2: Distribution of responses of all survey participants who reported participating in an on-line discussion on the 7-point Likert Scale when asked to rate the usefulness of the forum.

Table 4.3: Numbers and percentages of participants who participated in the On-line Discussion Forums and reported finding the sessions extremely useful, middle of the scale and not at all useful.

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th>Prep teachers</th>
<th>Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>% participated in on-line discussion forums</td>
<td>30</td>
<td>40</td>
<td>5</td>
</tr>
<tr>
<td>% found it extremely useful</td>
<td>3</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>% found it middle of scale</td>
<td>5</td>
<td>16.7</td>
<td>2</td>
</tr>
<tr>
<td>% found it not at all useful</td>
<td>2</td>
<td>6.7</td>
<td>-</td>
</tr>
</tbody>
</table>

The qualitative comments about the ongoing on-line discussion forum varied considerably. This seems to be at least in part related to teachers’ aptitude, prior experience, and prior attitude toward such forums. Many teachers commented that they read the material posted but did not offer their own posts, qualitative reports seem to suggest that those who did participate did so by asking questions, usually directed to QSA staff. As was the case with responses about the on-line training sessions there was a variety of opinions about the level at which the advice was pitched and its usefulness. There was little discussion by teachers, and some teachers commented on the dearth of
teacher participation. For some respondents the quantity of emails deposited in their inboxes was an issue, others reported having difficulty understanding how to participate at all.

CONSIDERATIONS:

Providing clear instructions of use using simple animations might be helpful to some having access difficulties.

Consideration of teacher expertise level – particularly for those who have already participated in using the materials and others who have not may be required in the first few years of implementation and beyond.

Look to other discussion forums for ways to both increase participation levels of participants, but also support the development of varied ways of participating beyond asking question of a forum leader, or being an observer.

QSA Staff Support

Forty-two percent of the Year 1 teacher participants (n = 26), twenty percent of Prep teacher participants (n=4), and sixty-three percent of Facilitator participants (n=14) reported having made contact with QSA staff to support them in the process of using Year 1 Checkpoint Assessments. Thirty-four percent of these participants (n = 15) found this support extremely useful, 5% (n = 2) found this support not at all useful and 9% (n = 4) identified with the middle point on this scale. Eighty-four percent of participants selected a response above midpoint when responding to this question. These results attest to the fact that many participants found the QSA staff involved in the trial organised, helpful and knowledgeable. Figure 4.3 represents these results in diagrammatic form while Table 4.4 provides percentages of respondents who found the additional support provided by the QSA team extremely useful, not useful at all and who responded at the midpoint of the 7 point Likert scale.
Figure 4.3 - Distribution of responses of all survey participants who reported contacting the QSA team for support on the 7-point Likert Scale when asked to rate the usefulness of the contact.

Table 4.4: Numbers and percentages of participants who contacted the QSA for advice or support and reported finding the contact extremely useful, middle of the scale and not at all useful.

<table>
<thead>
<tr>
<th>Year 1 Prep teachers</th>
<th>Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>% contacted QSA staff for support</td>
<td>26</td>
</tr>
<tr>
<td>% found it extremely useful</td>
<td>9</td>
</tr>
<tr>
<td>% found it middle of scale</td>
<td>2</td>
</tr>
<tr>
<td>% found it not at all useful</td>
<td>1</td>
</tr>
</tbody>
</table>

The qualitative comments left about the support of QSA staff comment positively on the speed of response, quality of replies, usefulness of support provided, and generally of the pleasant efficiency of the QSA staff involved. While two participants suggested that their enquiries were not adequately serviced (Didn’t really get my question answered), the majority of comments made, praised the staff involved. One respondent raised an important issue in this comment:

The people we dealt with were lovely. However they needed to be better informed by trial organisers about the contextual issues of some schools;
materials arrived two weeks later than other schools as no consideration was made for our limited mail service.

To be fair it is more likely that the tight timelines caused these issues with late materials, however the respondent raises important issues related to a centralised system in a state that services such wide regions across a large area. As such this issue is worth mention and consideration.

There were also calls for further opportunities to be involved in training with the QSA staff, and comments about their general professionalism:

  We were treated as professionals at all stages of the process.

**CONSIDERATIONS:**

The levels of expertise, professionalism, and ability to answer questions credibly have been high in the trial support team during the Supported Trial. This is no doubt related to the variety of expertise and experience within the team, and the professional approach taken by leaders and staff alike. The provision of this support was obviously important to the school-based staff involved and key to their attempts to participate in the Year 1 Checkpoints trial, and this should be a key consideration in any future implementation decisions.

**EQ Staff Support**

Thirty-four percent \((n = 21)\) of the year 1 teacher sample, twenty percent of the Prep teacher sample \((n = 4)\), and twenty three percent \((n = 5)\) reported having contacted other EQ staff – other teachers, HOCs, Administration Staff, colleagues and Key Teachers – to support them with using the Year 1 Checkpoint Assessments. Of these participants, 40% \((n = 12)\) found this support extremely useful, 17% \((n = 5)\) identified with the middle of the scale and 3% \((n = 1)\) found this support not at all useful. The most number of participants selected “Extremely Useful”. So the responses to this question provide us with a picture of teachers and administrators helping each other complete a task by networking and collaborative and collegial support.

There were comments made about the ineffectiveness of accessing teachers through busy administrators’ emails or other communication through Administrative Staff. Making an assumption that teachers have been informed because information has been sent to a school contact person who may or may not be involved proved problematic for some. This draws to the fore the necessity of a Key Teacher type role or for clearly set delineations of responsibility for existing Administration or Curriculum roles in the school.
Figure 4.4: Distribution of responses of all survey participants who reported contacting the other EQ staff members for support on the 7-point Likert Scale when asked to rate the usefulness of the contact.

Table 4.5: Numbers and percentages of participants who contacted other EQ staff for advice or support and reported finding the contact extremely useful, middle of the scale and not at all useful.

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th>Prep teachers</th>
<th>Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>% contacted EQ staff for support</td>
<td>21</td>
<td>34.4</td>
<td>4</td>
</tr>
<tr>
<td>% found it extremely useful</td>
<td>8</td>
<td>38.1</td>
<td>2</td>
</tr>
<tr>
<td>% found it middle of scale</td>
<td>4</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>% found it not at all useful</td>
<td>1</td>
<td>4.8</td>
<td></td>
</tr>
</tbody>
</table>

CONSIDERATION:
It is clear that participants in the trial benefited from the expertise of colleagues and peers. Are there opportunities for this expertise to be utilised more fully as part of the capacity building components of the Year 1 Checkpoints Assessment processes?

Lines of responsibility and delineation of responsibilities for the ‘process’ of the Year 1 Checkpoints suite need to be clearly organised and communicated to schools and teachers. Is there a role for a school-based organiser and trainer?
The timing of the project.
The timeline for the trial was truncated in its original design, and became more so as the trial continued. While these tight timelines were perhaps unavoidable, they have implications for the findings and recommendations made in this report, and for the reception of the materials more generally of the teachers involved.

Table 4.6 provides details of the initial advice provided to Supported Trial schools on expectations and requirements related to Supported Trial participation.

Table 4.6: Initial advice provided to Supported Trial schools on expectations and requirements related to Supported Trial participation.

<table>
<thead>
<tr>
<th>Month</th>
<th>Week beginning</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>July</td>
<td>26</td>
<td>EQ schools informed if accepted into supported trial</td>
</tr>
<tr>
<td>August</td>
<td>2 - 5</td>
<td>Nominated teacher participates in on-line information session. One computer connection per school.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Return completed Attachments B, C to QSA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Receive email regarding planning for regional cluster workshops. (You will need to confirm numbers attending.)</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Confirm school assessment selections</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Begin to trial June assessments: Either reading or writing and one numeracy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Return completed Attachment F: Workshop attendance confirmation by Tuesday 10 August 2010</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>Trial June assessments: Either reading or writing and one numeracy.</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>By 20th</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete Data Analysis Assessment Record (DAAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete External Evaluator June assessment survey</td>
</tr>
<tr>
<td>Aug-Sept</td>
<td>30</td>
<td>One day workshops</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monday 30 August: Workshop (North QLD - Cairns)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wednesday 1 September: Workshop (S.E. QLD - Brisbane)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Friday 3 September: Workshop (Central QLD - Rockhampton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trial one complete October assessment</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Continue trialling October assessment</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Continue trialling October assessment</td>
</tr>
<tr>
<td></td>
<td>20-27</td>
<td>School vacation</td>
</tr>
<tr>
<td>October</td>
<td>4</td>
<td>Continue trialling October assessment</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Continue trialling October assessment By 15th</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete Data Analysis Assessment Record (DAAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete External Evaluator October assessment survey</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Pupil Free Day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participate in Yr 1 Checkpoint Assessment moderation with other local schools in trial.</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>Trial February checkpoint assessment with Prep children</td>
</tr>
</tbody>
</table>
Data collected through survey responses, interviews and visits to schools and Supported Trial events provides insight into the impact of the short timelines on the results of the trials. This includes data about not having sufficient time between checkpoints, attempting to complete the Year 1 Checkpoints Assessments (October) and contextualising activities in addition to teaching, learning and assessment activities that had already been planned for the time period, and in some schools that already had school-based centralised assessment routines and systems teachers reported necessary repetition rather than overlap. With more notice of participation and the expectations of the Year 1 Checkpoints Assessments many of these concerns would have been addressed quite simply, however this context is an important consideration for the results of the Evaluation and the levels of uptake of the materials evident. An analysis of the qualitative data related to time provides several key messages.

It is clear that teachers were rushed and that workload was heavy because of the decision to trial the materials within a truncated timeline. Some teachers were able to consider their concerns about this truncated timeline, and their opinions of the Year 1 Checkpoints Assessments as separate. See for example these statements by respondents which demonstrate such an approach:

*I can see that given a more reasonable time frame that the information would be gathered with less stress...*

*I found it very difficult to implement in the timeframe, however I personally think they would be great given the opportunity to use them over a longer timeframe.*

Others were clearly not able to separate the implications of the truncated timeline from the assessment process itself. The volume of qualitative data extracts across a variety of sources that list timing issues related to the trial as a concern suggests that the truncated timeline affected the uptake of the Year 1 Checkpoints in many schools and classes. This has no doubt had an effect on teachers’ opinions of the materials themselves. It has also blurred the link between assessment and planning for some teachers as explained by this year 1 teacher:

*I just assessed them and there was no time to plan or teach in between.*
This led some teachers to consider ways to cut time, and made the link between assessment tasks as the basis for future planning and teaching more tenuous.

Additionally, many teachers made comments about time in relation to the Year 1 Checkpoints more generally.

*They were very time consuming to administer – this is a huge concern.*

Teachers expressed concern at the time individual one-to-one assessment tasks took with reference to how this could be organised for a full class. Other teachers were concerned with the number of assessment tasks. Some called for release time from teaching, others stated that without release time organised at their school context they would not have been able to complete the work. Some queried whether the assessment process should be structured across a longer time period, rather than expecting the assessments to be conducted ‘in-bulk’ as one teacher expressed it.

Issues of timing do need to be considered in any implementation plans for the Year 1 Checkpoints Assessments. However we caution against letting time issues drive all modifications of either the materials or expectations. It is clear that the materials must be usable across a variety of contexts, including multi-age contexts, and that time and workload must be considered. It is equally important to consider the principles of assessment as classroom practice and the framework of assessment for learning which is at the foundation of the Year 1 Checkpoints. Teachers should and must be supported without doubt in being able to reach the expectations, and provision of this support may be as important as adjusting the expectations themselves. It is important that the assessment and implementation plans are relevant to the real life worlds of Year 1 classrooms and that workload issues are considered. Time is a recurring theme in the data collected and will be dealt with throughout the analysis presented in this chapter.

**The Materials as a Diagnostic Tool**

As discussed in earlier sections of this report, when analysing an assessment tool it is important to make a distinction between “frameworks against which students’ learning might be reported and monitored, and vehicles through which evidence of student learning might be collected” (Forster, 2009, p. 15). The Year 1 Literacy and Numeracy Indicators are a framework. The Year 1 Checkpoints Assessments tasks are instruments – that could be deployed diagnostically - to allow teachers to collect information and evidence to subsequently make judgments about a student’s or group of students’ achievement, or otherwise, of a milestone or expectation – in this case a Year 1 Indicator. In this report we distinguish between the Framework and the Instruments so as to allow for more specific analyses and considerations than might be possible if the tool was discussed generally only.
The Framework - the Year 1 Literacy and Numeracy Indicators

The Indicators are end of year milestones or expectations that work somewhat as content standards. They describe the things that students should be able to do in literacy and numeracy by the end of Year 1. They provide an appropriate foundation for accountability purposes (Klenowski & Wyatt-Smith, 2010, p. 111) representing the level of achievement as a holistic judgment (yes or no). As such they differ from the notion of quality standards that teachers have become more accustomed to over the past decade’s standards based reform moves in Queensland, and to benchmark standards that describe a “typical outcome” (Maxwell, 2002, p. 2) usually with a definition of what constitutes a minimum acceptable level (Klenowski & Wyatt-Smith, 2010). As a set of P-9 Literacy and Numeracy Indicators the suite works as a set of progressive targets or expectations (Maxwell, 2002, 2008) for end of year junctures. They focus the work of teachers and students toward achieving outcomes in the two cross curricular components, literacy and numeracy, by providing a framework to make judgments over time of what it means to learn to be literate and numerate (Klenowski & Wyatt-Smith, 2010) in Queensland schools. While their actual content is similar in both quality and format to other taxonomies of the developmental pathway, such as the Year 2 Diagnostic Net continua for example, because they frame the ‘indicators’ as end of year expectations of the learning that will have occurred, and not as assumed mono cultural development milestones, they could lead to a focus on planning and teaching rather than individual or cohort student capacities if used and talked about in particular ways. This is a tenuous expectation that would rely heavily on the professional knowledge of teachers and the expectations for use placed on the Indicators.

While we agree with Forster (2009, p.14) that “the diagnostic instrument(s) is central to the broad set of materials described... as diagnostic tools”, there are several considerations related to using the Indicators as the framework for this early years diagnostic tool that warrant mention here before moving on to discuss the actual checkpoints assessments as diagnostic instruments.

Forster (2009) uses a notion of diagnostic power, which relates to the potential of the diagnostic tool to uncover, and make visible specific strengths and weaknesses in a student’s abilities, capabilities and skills. By its nature a framework can only ever describe levels of achievement, or in the case of the Year 1 Indicators describe actual capabilities or expected learning - things that have been learnt or can be done. They cannot be used to collect evidence of these capabilities nor to expose specific strengths and weaknesses of a students’ performance, although they can map these strengths and weaknesses. So their diagnostic power is relatively weak. However, as the foundation of a tool, the importance of the scope of the framework, and its ability to be fair, reliable and valid is relevant for consideration.

This study was not required to address questions about the Year 1 Indicators as part of its brief. However we make these considered comments as feed
forward for the team tasked with taking the Year 1 Checkpoints forward. These include suggestions that the team:

- Consider if the Year 1 Indicators accurately describe the Literacy and Numeracy accomplishments considered appropriate for all Queensland students at the end of Year 1. This is particularly important as Australian education expectations are adjusted to take account of the Australian Curriculum. However regardless of the final form or timeline of this new curriculum document the Indicators must stand up to rigorous analysis based on the theoretical underpinnings of the models of literacy and numeracy, language and learning that are in play. These models need to be clearly stated and justified and not just assumed. The foundations of the Indicators and the Checkpoints need to be aligned, and as the Checkpoints currently do not make any statement about the understandings of literacy and numeracy at their foundation it is difficult to ascertain if this is the case. Are the Indicators founded in the best current understandings of literacy and numeracy, and literacy and numeracy teaching?

- Consider whether the Year 1 Indicators record information on the achievement of skills that actually underpin what it means to be literate or numerate in school and beyond. The Indicators are not capturing in any way the full scope of literacy and numeracy capabilities of students. They may be able to capture whether students have been taught and have learned particular school literacy and numeracy skills that are valued within our current context. Is there a clear understanding of what the Indicators can record, or are they assumed to represent ‘literacy’ and ‘numeracy’ in some absolute sense?

- Consider the Indicators’ capacity to provide a fair and accurate description of literacy and numeracy capabilities for all students regardless of their language, cognitive, social, cultural and physical needs. They should provide specific and useful information to teachers about how to use the Indicators to plan, teach and assess, in an inclusive manner. Are the Indicators fair and just to a diverse student cohort?

- Consider the Year 1 Indicators in relation to their capacity to take account of cultural and language differences for those students who are likely to be disadvantaged by assumptions and bias toward mainstream, middle class understandings of literacy and numeracy. Are the Indicators fair?

- Consider issues of content validity in relation to the Year 1 Indicators. Do the Indicators adequately describe the scope of literacy and numeracy competence in a broad and balanced sense. Their current focus on Speaking and Listening, Reading and Viewing, and Writing and Designing in literacy and Number, Algebra, Measurement, Space, and
Chance and Data in numeracy suggests a balanced and broad understanding of literacy and numeracy. However it is important to ensure that the Indicators within each domain adequately describe the domain, and that the composite of Indicators adequately describes a broad and balanced notion of literacy and numeracy. Are the Indicators valid and reliable?

- Further consider issues of content validity as any adjustments to the Indicators are made to fit with new curriculum documents. Currently the Indicators claim to be linked to a broad range of state based curriculum documents and support materials, however as shifts in expectations are made to fit the Australian Curriculum, ensuring that the Indicators remain linked to a range of other important documents will be fore grounded. Are the indicators linked to a broad range of teacher support and curriculum resources or narrowly defined through a connection with one document?

Within the qualitative data there were several participants who asked for clarification about the origins, foundations, format and structure of the Year 1 Indicators.

There needs to be much more information about the indicators if they are going to replace the Net. They should be teased out with dot points below each indicator. There is too much room for personal judgment and differences in the Indicator.

In interviews and conversations teachers discussed not having used the Indicators in their planning or assessment prior to becoming involved in the Trial. Some discussed that they had not known of the Indicators prior to this time. What this lack of knowledge may lead to is a misrepresentation of the Indicators, their purpose and capacity. The legacy of teachers working with the Year 2 Net continua since the nineties is a constraint of teachers’ understandings of literacy development through the lens of normative developmental pathways, somehow linked to the capacity of individuals. Once the Year 1 Indicators become represented as a developmental pathway, and not as end of year expectations of teaching and learning, the Indicators will be used to talk about children, individual capacities and development. This distinction seems small – but it is an important one in relation to ensuring a focus on assessment to inform teaching and learning.

CONSIDERATIONS

The theoretical foundations, assumptions and links to other assessment practices of the Year 1 Indicators needs to be clear and explicit to teachers expected to utilise the Indicators.
Training for further iterations of the Year 1 Checkpoint Assessments must focus on the Framework (the Year 1 Indicators) of the tool as well as the instruments, and developing teachers’ understandings of the Indicators, their relationship to standards and expectations in other areas and assessment processes, and their position within the P-9 Indicators more generally.

The Indicators must be included in any Year 1 Checkpoints package. The alignment of the framework and the tools must be explicit for those expected to utilise the Checkpoints.

A review of the content of the Literacy and Numeracy Indicators as they relate to current understandings of literacy and numeracy, and current and future curriculum expectations of literacy and numeracy across the curriculum should be undertaken.

The justification used to prioritise the selection of targeted P-1 Indicators within the Year 1 Checkpoint Assessments needs to be clearly stated within the support materials that wrap around the Year 1 Checkpoints materials.

The Instruments – The Year 1 Checkpoints Assessments

The analysis of the Year 1 Checkpoint Assessments and materials was conducted in three stages:

- The review of the instruments\(^4\);
- The use of surveys to collect information from participants in the Supported Trial, and the descriptive quantitative analysis of this data for trends; and
- The coding of qualitative data collected via surveys and interviews, and work samples collected from participants to exemplify the trends identified.

**Review of the Instruments**: The Instruments were reviewed calling on a framework put forward by Forster (2009). We considered the assessment tools in terms of:

1. Their fit for purpose and technical issues related to using the tools in particular contexts;
2. The match between the intent of the tool and what the tool can actually achieve;
3. The clarity and explicitness of materials provided to support teachers in understanding what is being assessed;
4. The inclusion of quality, clear directions and instructions to teachers for implementation and next steps;

\(^4\) This review was supported by the expert comments of Karen Dooley (QUT) and Beryl Exley (QUT). Their input to the analysis of the materials presented in this chapter is acknowledged here.
5. The quality of advice given about the needs of specific students and implementation issues related to this;
6. The level to which the instruments would engage the interests of the target group of students;
7. Their validity (their ability to capture evidence of students’ literacy and numeracy – do they assess what they say they assess?); and
8. Their reliability (their ability to consistently and accurately capture information on students’ performance in literacy and numeracy – are the assessments fair?).

**On-line Surveys:** Three groups were surveyed and asked to respond to questions that related to the components of the Year 1 Checkpoints process that they had been involved in. All respondents were asked about the Checkpoints generally and the processes of assessment in the early years of schooling. Year 1 teachers were also asked to comment on the June and October materials, Prep teachers on the February materials, and Facilitators asked to comment on the processes of implementing all three Checkpoints Assessments at their school. Table 4.7 details the numbers of respondents to each of the surveys and their particular school based roles. The bottom line of the table provides the numbers of respondents to each survey who actually responded to the survey questions. Other categories of respondents for whom the questions were not relevant but who entered the survey were sent on a pathway to the end of the survey and are thus not included in any analysis of data.
Table 4.7: Number of respondents to each of three surveys and the school based roles of these respondents.

<table>
<thead>
<tr>
<th>ROLE</th>
<th>Year 1 Teachers' Survey</th>
<th>Prep Teachers' Survey</th>
<th>Facilitators' Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Head of Curriculum</td>
<td>4</td>
<td>5.3</td>
<td>-</td>
</tr>
<tr>
<td>Prep Teacher</td>
<td>6</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Prep/Year 1 Teacher</td>
<td>7</td>
<td>9.3</td>
<td>3</td>
</tr>
<tr>
<td>Year 1 Teacher</td>
<td>43</td>
<td>57.3</td>
<td>3</td>
</tr>
<tr>
<td>Year 1/Year 2 Teacher or other combination</td>
<td>11</td>
<td>14.7</td>
<td>-</td>
</tr>
<tr>
<td>Deputy Principal</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>5.3</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>28</td>
<td>22</td>
</tr>
</tbody>
</table>

The sample size of respondents to all surveys, but particularly the Prep teachers’ and Facilitators’ surveys, was small. This was primarily a result of the timing and timelines of the Supported Trial. The small sample size has meant that analysis has had to remain descriptive.

The respondents came from a variety of school sites and represent a sample across the key school type characteristics identified in the original sample selection processes. That is: schools from three regions; those with high Indigenous student populations; those from rural, remote, provincial and metropolitan locations; those with single and multi-age classes; and those with a high proportion of students from low SES backgrounds. Table 4.8 shows the distribution of respondents across these categories. The number in the brackets in each cell represents the number of school sites representing each category in the full Supported Trial sample schools.
Table 4.8: Distribution of respondents across priority school categories n (number of participants in sample in each school category)

<table>
<thead>
<tr>
<th>SES Status</th>
<th>Low SES</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yr 1 - 4</td>
<td>Yr 1 - 57</td>
</tr>
<tr>
<td></td>
<td>Prep - 2</td>
<td>Prep - 25</td>
</tr>
<tr>
<td></td>
<td>Fac - 4</td>
<td>Fac - 24</td>
</tr>
<tr>
<td></td>
<td>(9 Schools)</td>
<td>(90 schools)</td>
</tr>
<tr>
<td>% Indigenous students</td>
<td>% Indigenous Students &gt; 10%</td>
<td>&lt; 10%</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 22</td>
<td>Yr 1 - 39</td>
</tr>
<tr>
<td></td>
<td>Prep - 5</td>
<td>Prep - 22</td>
</tr>
<tr>
<td></td>
<td>Fac - 8</td>
<td>Fac - 20</td>
</tr>
<tr>
<td></td>
<td>(26 Schools)</td>
<td>(73 schools)</td>
</tr>
<tr>
<td>Size of school</td>
<td>Small</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 6</td>
<td>Yr 1 - 37</td>
</tr>
<tr>
<td></td>
<td>Prep - 5</td>
<td>Prep - 15</td>
</tr>
<tr>
<td></td>
<td>Prep - 6</td>
<td>Fac - 16</td>
</tr>
<tr>
<td></td>
<td>(18 schools)</td>
<td>(58 schools)</td>
</tr>
<tr>
<td></td>
<td>% Indigenous Students</td>
<td>% Indigenous Students</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 19</td>
<td>Yr 1 - 18</td>
</tr>
<tr>
<td></td>
<td>Prep - 15</td>
<td>Prep - 7</td>
</tr>
<tr>
<td></td>
<td>Fac - 11</td>
<td>Fac - 6</td>
</tr>
<tr>
<td></td>
<td>(20 Schools)</td>
<td>(35 schools)</td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Region</td>
</tr>
<tr>
<td></td>
<td>Metropolitan</td>
<td>Province</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 29</td>
<td>Yr 1 - 8</td>
</tr>
<tr>
<td></td>
<td>Prep - 15</td>
<td>Prep - 1</td>
</tr>
<tr>
<td></td>
<td>Fac - 12</td>
<td>Fac - 4</td>
</tr>
<tr>
<td></td>
<td>(38 schools)</td>
<td>(13 schools)</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>Remote</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 18</td>
<td>Yr 1 - 6</td>
</tr>
<tr>
<td></td>
<td>Prep - 7</td>
<td>Prep - 4</td>
</tr>
<tr>
<td></td>
<td>Fac - 6</td>
<td>Fac - 8</td>
</tr>
<tr>
<td></td>
<td>(35 schools)</td>
<td>(26 schools)</td>
</tr>
<tr>
<td></td>
<td>Region</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South Qld</td>
<td>Central Qld</td>
</tr>
<tr>
<td></td>
<td>Yr 1 - 19</td>
<td>Yr 1 - 20</td>
</tr>
<tr>
<td></td>
<td>Prep - 15</td>
<td>Prep - 8</td>
</tr>
<tr>
<td></td>
<td>Fac - 11</td>
<td>Fac - 9</td>
</tr>
<tr>
<td></td>
<td>(20 Schools)</td>
<td>(35 schools)</td>
</tr>
</tbody>
</table>

Opinions about the Year 1 Checkpoint Assessment Materials Generally
The Year 1 Checkpoint Assessments aim to implement an approach to assessment that shifts the focus of assessment toward informing future teaching for all students in Year 1, rather than identifying students who are not achieving particular indicators for additional intervention. This purpose in and of itself is educationally useful. It situates assessment within the broader professional pedagogical and curriculum practice of teachers, and foregrounds teaching as implicated in student learning. With the provision of explicit training and advice for teachers it has the potential to be part of an approach toward building effective assessment practice within the early years.

Despite the fact that there are advantages to the approach taken there remain issues of concern with the Year 1 Checkpoints Assessments materials generally. These concerns would be addressed in some ways at least by the
inclusion of succinct, focused, materials that front load information about the Year 1 Checkpoints as a systems-based approach to early years’ assessment.

Our analysis is supported by the quantitative and qualitative data collected from teachers who have been involved in the Supported Trial during 2010. The remaining general concerns include:

**Issues related to the clarity and explicitness of links to purpose and understandings of literacy and numeracy**

- There is currently no clear statement of purpose of the Year 1 Checkpoints that unpacks the links between the Assessments and the framework used to record information collected (the Year 1 Indicators);

- Nor is there a clear and explicit statement of the understandings of literacy and numeracy being called upon within the Assessment materials. This leaves teachers with no statement about what the assessments are assessing (What model of literacy and numeracy is being deployed here? What components of broader concepts of literacy and numeracy are being prioritised for assessment and why?) or a clear understanding of how the Year 1 Checkpoints fit with other tools and initiatives utilised in Queensland schools – particularly in the first few years of schooling.

Currently without such front loading statements as above, the validity of the Year 1 Checkpoints as a suite is difficult to comment upon because what is being assessed has been left implicit.

While the ‘future directions’ focus in the June assessments provides information about where to next for teachers, the tools themselves focus on what students can do rather than what they cannot do. Thus the diagnostic possibilities of the assessments rely wholly on teachers’ capacity to use the evidence collected in diagnostic ways for individuals, groups and the class cohort. While we do not suggest a shift in this focus necessarily, we do recommend that there is a strong focus on the provision of specific advice about the processes of using evidence collected to inform future teaching and learning. This should be seen as a capacity building activity.

In addition there remain issues related to the potential of the assessments to be deployed in fair, equitable, valid and reliable ways. These issues would in part be addressed by the development of a larger set of assessments – that work equivalently to assess achievement of particular Indicators.

**Issues related to the culture fairness of the assessments;**

- There is currently little recognition of the content validity issues that arise for assessment if items, tasks and instruments do not take account, in educationally useful ways, of the diversity of experience, language use, understandings and difference that a diverse cohort of students bring to schooling.
Issues related to the differentiation of assessment necessary to provide inclusive assessment processes for all students;

- There is currently little information (except for that information included within the Implementation Advice of the October Assessments) provided for teachers about appropriate and inclusive ways to differentiate assessment to ensure an equitable and inclusive approach to assessment for those students who may have specific cognitive, social, language, behavioural, physical or cultural needs for learning.

Responses to the surveys indicate that many teachers adjusted the assessments to fit specific needs of a diverse range of students. Thirty five percent of both the Year 1 and Prep teachers indicated that they had adjusted the assessment processes or expectations for a particular student or students in their class. Facilitators (63%) suggested that many teachers at their schools had also made such amendments. The variety of students that teachers and facilitators listed as requiring these amendments included those with specific physical, language, cognitive, social, behavioural and cultural needs. When asked to explain why they had felt the need to make adjustments teachers described principles of inclusion and support as important, along with issues raised about the educational imperative of students achieving success. With no support or provision for advice about how to adjust or amend the tools, the processes or the expectations of the Year 1 Checkpoints, teachers were left to design these links themselves. It is likely that there were differential effects in how these processes of differentiation played out and were recorded.

Participants were asked to provide their level of agreement (strongly agree to strongly disagree) to ten positive statements about the Year 1 Checkpoints materials. In all cases the response most likely to be selected was above a code of 5 (a positive response), with low numbers of respondents strongly disagreeing with any of the statements. Table 4.9 shows the distribution of responses across the 7 point Likert scale on the ten statements used to query respondents’ opinions about the materials generally.
Table 4.9: Distribution of Responses across the 7-point Likert scale on level of agreement to statements about the materials generally. (Shading represents the most popular response for each statement)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment materials would support me in quality assessment for year 1 students.</td>
<td>5</td>
<td>4.9</td>
<td>5</td>
<td>4.9</td>
<td>7</td>
<td>6.8</td>
<td>10</td>
</tr>
<tr>
<td>I found the Year 1 Literacy and Numeracy Checkpoint Assessment materials easy to access on-line.</td>
<td>3</td>
<td>2.9</td>
<td>2</td>
<td>1.9</td>
<td>7</td>
<td>6.8</td>
<td>8</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment materials would support me to differentiate assessment so that I meet the needs of all of my students and their diverse needs.</td>
<td>9</td>
<td>8.7</td>
<td>7</td>
<td>6.8</td>
<td>7</td>
<td>6.8</td>
<td>14</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment materials would help to inform teaching and learning in my class.</td>
<td>3</td>
<td>2.9</td>
<td>5</td>
<td>4.9</td>
<td>4</td>
<td>4.9</td>
<td>8</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment materials would support me to differentiate pedagogy so that I meet the diverse needs of all of my students.</td>
<td>4</td>
<td>3.9</td>
<td>8</td>
<td>7.8</td>
<td>6</td>
<td>5.8</td>
<td>18</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment work samples would provide support for me as I judged my students' work and the effectiveness of my teaching.</td>
<td>3</td>
<td>2.9</td>
<td>5</td>
<td>4.9</td>
<td>4</td>
<td>3.9</td>
<td>11</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment work samples would provide support for me as I planned subsequent teaching and learning sequences.</td>
<td>3</td>
<td>2.9</td>
<td>4</td>
<td>3.9</td>
<td>3</td>
<td>2.9</td>
<td>9</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment materials complement my current approach to assessment in year 1.</td>
<td>4</td>
<td>3.9</td>
<td>5</td>
<td>4.9</td>
<td>1</td>
<td>1</td>
<td>10.7</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment Data Analysis Assessment Record (DAAR) provides a useful way to record class results on the Checkpoint Assessments.</td>
<td>5</td>
<td>4.9</td>
<td>4</td>
<td>3.9</td>
<td>1</td>
<td>0</td>
<td>9.8</td>
</tr>
<tr>
<td>The Year 1 Literacy and Numeracy Checkpoint Assessment Data Analysis Assessment (DAAR) Record would provide useful support for me as I planned subsequent teaching and learning sequences.</td>
<td>5</td>
<td>4.9</td>
<td>6</td>
<td>5.9</td>
<td>3</td>
<td>2.9</td>
<td>7</td>
</tr>
</tbody>
</table>
Respondents were also asked to comment on their level of agreement to statements that related to the theoretical foundations of the materials and processes. Table 4.10 summarises the distribution of responses to these 6 statements. Again the responses are generally positive with 5 or 6 being the most selected code in all cases.

**Table 4.10: Distribution of responses across a 7 point Likert scale when year 1 teachers were asked to comment on their level of agreement (strongly disagree to strongly agree) to statement about the foundations of the Year 1 Checkpoints.**

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The theoretical framework of assessment for learning is effective for a systematic assessment process in the early years.</td>
<td>1 1.7</td>
<td>2</td>
<td>3.3</td>
<td>7</td>
<td>11.7</td>
<td>20</td>
<td>33.3</td>
</tr>
<tr>
<td>The Year 1 Checkpoint Assessment process would support capacity building of teacher assessment expertise at my school.</td>
<td>3 4.9</td>
<td>2 3.3</td>
<td>3 4.9</td>
<td>2 3.3</td>
<td>27 44.3</td>
<td>15</td>
<td>24.6</td>
</tr>
<tr>
<td>The Year 1 Checkpoint Assessment process encouraged collaborative planning across teachers at my school.</td>
<td>6 10</td>
<td>7 11.7</td>
<td>5 8.3</td>
<td>5 8.3</td>
<td>13 21.7</td>
<td>17</td>
<td>28.3</td>
</tr>
<tr>
<td>The Year 1 Checkpoint Assessment process provided me with new information about the achievement levels of my students.</td>
<td>5 8.2</td>
<td>2 3.3</td>
<td>3 4.9</td>
<td>11 18</td>
<td>15 24.6</td>
<td>13</td>
<td>21.3</td>
</tr>
<tr>
<td>The Year 1 Checkpoint Assessment process supported me in providing access to quality instruction for all my students.</td>
<td>6 9.8</td>
<td>1 1.6</td>
<td>4 6.6</td>
<td>13 21.3</td>
<td>16 26.2</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>The Year 1 Checkpoint Assessment process supported me to learn more about assessment in the early years.</td>
<td>6 9.8</td>
<td>2 3.3</td>
<td>6 9.8</td>
<td>7 11.5</td>
<td>13 21.3</td>
<td>19</td>
<td>31.1</td>
</tr>
</tbody>
</table>

While the descriptive representation above indicates that teachers and facilitators who replied to the surveys generally agreed that the Year 1 Checkpoints Assessments were useful assessment practice for Year 1 in Queensland, the participants in the trial had issues to raise about these materials also. Along with specific item by item based feedback, which has been passed onto the QSA in other forms, teachers and other school staff raised issues related to:
• access to materials such as reading texts;
• issues related to access to teacher relief;
• issues related to the over reliance on one-to-one assessment techniques and the implications of this for assessing all students; and
• timing and workload.

In the sections that follow, we address each of the Checkpoints to discuss their use as assessment instruments in Year 1 and to provide feed forward suggestions to those tasked with the continuing development of the Year 1 Checkpoints.

Assessments implemented in February Assessment Checkpoint

The February Checkpoints Assessment materials were implemented during November, with Prep students and teachers in the 2010 Supported Trial. In many cases Prep teachers at schools were not informed about the expectation for them to participate until late in the year. This led to a situation where they were left with little time to become familiar with the materials, and had a short time period to complete the assessments also. The implications of these decisions on any data collected about the February assessments are twofold. Firstly the fact that teachers were collecting information about students at the end of the year when there would be no opportunity to use the data to inform future teaching has meant that teachers have considered the role of the assessments as being summative and not formative. Secondly, the organisational issues related to timing could certainly be seen as implicated in teacher uptake of the materials and their focused practices. The sample size of respondents who replied to the survey is very small. This no doubt is related to issues of timing of the Assessments implementation.

The February Checkpoints target students’ achievement of selected Prep indicators by suggesting everyday learning and teaching activities that provide opportunities for evidence to be collected and subsequent judgments to be made about students’ achievement of end of Prep year expectations in literacy and numeracy. The activities themselves are useful early years’ activities for learning and teaching components of literacy and numeracy. As suggestions for assessment contexts and processes, the February Checkpoints provide the opportunity for teachers to ensure that students are assessed in culturally relevant, reliable and equitable ways for the contexts in which they teach. However without advice provided to teachers about these important issues, the implementation of the Assessments in fair and equitable ways is likely to be variable.

Despite the tight timelines that they worked under, many of the small sample of teachers who responded to the survey reported finding the experience useful, and the general consensus from those who replied to the survey was that the materials were effective early years assessment processes (see table 4.11).
Table 4.11: Distribution of Responses across the 7-point Likert scale on the extent to which the February assessments were judged as effective by Prep teachers. (Shading represents the most popular response for each statement). Note the very small sample size.

<table>
<thead>
<tr>
<th>Area</th>
<th>Not at all effective</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Extremely effective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Speaking and Listening</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Reading and Viewing</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Writing and designing</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Number</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Algebra</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Measurement Space</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

While the small numbers of respondents make interpretation from Table 4.11 above difficult, the data does indicate that the Prep teachers who replied to provide information about the February checkpoints considered the February Assessments to be effective early years assessment. The qualitative comments provide further evidence for this claim. For although teachers did comment on time constraints, and had modification advice for particular domain components, the general tone of the qualitative comments were positive. Only one negative comment was collected from the survey data when teachers were asked for further information about the effectiveness of the February Checkpoints Assessment.

So the February Assessments provide advice about effective contexts and processes for assessing young children. However in their current form there are several issues that require addressing, if they are to be considered as more than general advice to teachers about quality assessment contexts for the early years of teaching. To begin with, the constructs of literacy and numeracy being utilised in the February Checkpoints are implicit at best. As with the materials generally, this is in part as a result of there being no statement of the theoretical understandings of literacy and numeracy at the foundation of the materials. In the case of the February Checkpoints however, the absence of any details of sequence of learning, implementation plan or future directions means that these understandings are left even more implicit than in the other Checkpoints. The work samples go some way to alleviate this, but teachers are left to understand from these samples what is being valued as literacy and numeracy, and this has grave consequences in terms of one document driving teachers’ understandings of literacy and numeracy and student learning. The Year 2 Diagnostic Net as a process has been critiqued for ‘becoming the curriculum’ (van Kraayenoord, Luke, Elkins and Land,
1999), but the consequences of exemplar work samples taking up this role is of even greater concern.

Secondly, the activities and contexts listed within the February Checkpoints are everyday classroom activities, the suggestion being that the products and processes of these activities that would be going on in classrooms regardless, be collected and then used to make judgments about students performance. This is well within the principles of effective early years’ assessment. To achieve the status of a point-in-time check – which they currently claim to be - the February Checkpoints need to be more than a list of advice on effective assessment contexts linked to particular Indicators. There would be a need for a broader range of support materials to be provided if the February Checkpoints are to be deployed in any capacity beyond a statement of advice about contexts to collect assessment data.

Thirdly there is no information for teachers about the diversity of the student cohorts and the implications of this diversity for collecting evidence of, analysing and recording student achievement. (This is further complicated because the evidence is being collected of achievement of end of year expectations, for a previous year of schooling that is not currently compulsory for all children). There is no advice to teachers about the implications for students who are not able to demonstrate any or most of the Indicators. As such, the usefulness of the assessment activities and contexts is likely to be variable for these students. This is likely to most disadvantage those students who are already at risk of being disadvantaged by schooling. This includes those who are using English as a second or additional language, those who have specific learning, behavioural, language, cultural or social needs that make engaging with school-based activities difficult in the first few weeks of school, and those who have specific learning needs. For these students the reliability of these assessments may be questionable, and teachers require training and advice to deal with this.

Finally there would seem to be an overuse of transcription data as evidence of students’ achievement of particular Indicators. The practicalities of this as a way to record evidence for cohorts of students, especially in the first few weeks of Year 1, need to be considered. Transcript data would usually be used for specific and targeted purposes and consideration of other means to collect evidence may be useful in the future development of the February Checkpoints.

The February Assessments currently act as solid advice to teachers about everyday classroom activities that can act as contexts and purposes for effective early year’s assessment. They are an adequate and useful list of activities and contexts for assessment of students’ literacy and numeracy achievement.

For the February Assessments to take a role beyond this there would be a need for the provision of a strong statement about literacy, numeracy, student diversity, language and school entry expectations; implementation advice;
support for planning future directions for students across a full range of achievement levels; advice on recording, analysing and reporting the results of student assessment; and specific advice on how this information can and should be used to inform planning for the first few weeks and months of Year 1, the February Checkpoints would be provide effective advice for year 1 teachers about the on-going assessment of students against the P-3 Indicators.

CONSIDERATIONS

There is a need to provide teachers with support materials that are wrapped around these suggested activities and contexts for assessment. If the February Checkpoints are to work as a tool to build assessment capacity in the early years then a commitment to providing these supports is vital. At the moment the validity, reliability and educational usefulness of the February Checkpoints as tools is highly dependent on the expertise, knowledge and content area understanding of the teacher involved. The purposes and intents that the information collected is put toward also have the potential to be variable.

The provision of training and advice to teachers about the adjustment or differentiation of assessment contexts and processes so that all students experience fair, reliable and equitable assessment should be of high priority.

Prep teachers were also asked to comment on the placement of the February Checkpoint Assessments and the logic of placing them in November in Prep or February in Year 1. There was support for both leaving them in Prep, with some respondents believing this would be a more appropriate in terms of the reliability of the evidence collected, and placing them early in Year 1 so that the link between assessment and future teaching and learning be highlighted. The actual assessment contexts of the February Assessments were well received by most Prep teachers who responded to the surveys. Many saw the benefit of end of year assessments, the results of which could be passed to Year 1 teachers along with data collected in the Early Learning Framework. The positive by-product of having a formal means of making and recording teacher judgments about student performance as the basis of shared conversations between Prep and Year 1 teachers about children, their capacities and about teaching and learning programs is evident and was raised by teachers in the study. However the purpose of the February Checkpoints is to provide information to year 1 teachers as they plan for literacy and numeracy teaching and learning with their students. For this purpose the placement of the February Checkpoints in February is optimum. The fact that Prep teachers were open to engaging in assessment as a somewhat summative measure of whether students had reached the end of Prep year expectations as represented by the Indicators supports the call for
the Year 1 Checkpoints to be considered as part of a systems-based approach to assessment across P-3.

CONSIDERATIONS

We are not suggesting the implementation of a Prep Checkpoint. However as part of a systems-based approach to assessment in the early years of school advice and suggestions of assessment contexts could be provided to Prep teachers, or early years teachers more generally. The decision on how this support could be used should then be a school-based decision.

The absence of any implementation advice, strategies, links to teaching strategies or support to teachers in using the evidence collected to inform teaching means that the diagnostic power of the February Checkpoints relies on pre-existing capacities of teachers in an absolute sense. The Year 1 Checkpoints Assessments aim to support and develop teacher capacity in assessment in the early years of schooling, and develop effective practices of assessment for learning – that is for using assessment to inform future teaching and learning – but this would not be achieved through the February Checkpoints without the provision of relevant support materials and focus.

Assessments implemented in June Assessment Checkpoint

The June Assessments Checkpoints were implemented in August during the 2010 Supported Trial. This had implications for teachers’ being able to capitalise on the information collected as a means to inform future teaching before the October Assessments were conducted. Time issues again impacted on the data that could be collected in the Evaluation, and the general uptake of the materials for teachers. There is evidence in the data that for some teachers at least the purpose of the June Checkpoint was somewhat less clear than we might have expected it to be if there had been adequate time to use the information before the next assessments were conducted. This has implications for the data collected about the instruments themselves.

The June Assessments Checkpoints include assessments designed to assess students’ achievement of selected Year 1 Indicators in literacy and numeracy. The Literacy suite includes reading/viewing and writing/designing tasks that are embedded within the context of ‘The Diverse Family’, along with a Letter Sort and Spelling Search task. The numeracy suite includes measurement and number tasks linked to the integrating device of ‘The Post Office’, along with a space task, ‘Round the Ridges’, and an algebra task, ‘Patterns in Songs and Games’. As with all of the Checkpoints the instruments do not attempt to assess all of the indicators, but instead target particular Indicators on which to focus.

In the Supported Trial, teachers had the opportunity to select a smaller number of assessments to implement from the full suite of June Checkpoints. This means that many teachers did not have the same workload as teachers implementing the June Checkpoints as part of the complete Checkpoints
process would have. *This brings to light implications for workload and time for teachers involved in the Year 1 Checkpoints in any future deployment of the materials.* The teachers in the Evaluation expressed much concern about the time taken to complete the February, June and October Checkpoints, but the June Checkpoints Assessments particularly. If in future implementations of the Checkpoints the task is even larger than it was in the 2010 Supported Trial, it is likely to require a great deal of implementation support.

To provide details of the patterns of participation of Year 1 teachers in the June assessments, teachers were asked in the survey to record which of the Checkpoints they engaged with. Sixty-one percent (n = 37) of participants who responded to the survey implemented the Reading and Viewing assessment and 57% (n = 35) implemented the Writing and Designing Assessment which included in this case the ‘Spelling Sort’ and ‘Letter Search’ tasks. The Reading and Viewing and Writing and Designing tasks were a component of ‘The Diverse Family’ Assessment. This means that at least some teachers decided to implement more of the assessment processes than was required of them in the 2010 Supported Trial, as teachers were asked to select either the Reading and Viewing or the Writing and Designing components. In comparison, the Numeracy Assessments allowed more choice for participants and this is represented in the assessment choices made by responding teachers. Twenty eight percent (n = 17) of participants implemented ‘Post Office – Measurement’ and forty-nine percent (n = 30) ‘Post Office – Number’, the highest number for any numeracy based checkpoint assessment. Those who implemented ‘Patterns in Songs and Games’ comprised 21% of the sample while 23 % implemented the ‘Round the Ridges’ Checkpoint. Table 4.12 details the assessment choices for the June Checkpoint made by year 1 teachers who responded to the Year 1 teachers’ Survey.

**Table 4.12: Assessment choices of Year 1 Teachers in June Checkpoint (Teachers may select multiple assessments)**

<table>
<thead>
<tr>
<th>Assessment Choices</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literacy Checkpoint Assessments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% implemented Reading and Viewing Assessment – The Diverse Family</td>
<td>37</td>
<td>60.7</td>
</tr>
<tr>
<td>% implemented Writing and Designing Assessment – The Diverse Family &amp; Spelling Search and Letter Sort</td>
<td>35</td>
<td>57.4</td>
</tr>
<tr>
<td>Numeracy Checkpoint Assessments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% implemented Patterns in Songs and Games – Algebra</td>
<td>13</td>
<td>21.3</td>
</tr>
<tr>
<td>% implemented Post Office – Measurement</td>
<td>17</td>
<td>27.9</td>
</tr>
<tr>
<td>% implemented Post Office – Number</td>
<td>30</td>
<td>49.2</td>
</tr>
<tr>
<td>% implemented Round the Ridges – Spaces</td>
<td>14</td>
<td>23</td>
</tr>
</tbody>
</table>

Qualitative comments from teachers in the field can be coded as relating to requests for changes to:
• Correct errors in the assessment instruments;
• Address difficulties with understanding the purposes of individual assessment requirements;
• Address difficulties with the format or structure of specific assessments;
• Address issues related to making amendments to the assessments for specific students or cohorts of students;
• Address issues related to the quantity of assessment items, and the format of different assessment items as they relate to issues of timing; and
• Address issues of time and workload more generally.

Respondents were also asked to rate the level of usefulness of the different components of each of the June Checkpoints that they had implemented with their students, across a 7 point Likert scale (not at all useful to extremely useful). The responses across all assessments and across all components were generally positive, although there was a full range of responses. The usefulness of the time allocation component of the Assessments was generally scored low. This, taken in combination with the qualitative responses collected, would lead to the verdict that teachers felt very strongly that the time allocation advice was underestimated and thus misleading.

As an example of the responses from teachers when asked about the usefulness of the components of each of the June Assessments table 4.13 demonstrates the distribution of responses across the 7 point Likert scale for *The Diverse Family*. This is used as a representative example of the distribution of responses for components of the June Assessments.
Table 4.13: Distribution of responses (n & %) across a 7 point Likert scale when asked to comment on the usefulness of components of each Assessment in the Diverse Family June Checkpoint. (Shading represents the most popular response for each statement)

<table>
<thead>
<tr>
<th>Assessment Component</th>
<th>Not at all useful</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Extremely useful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Context</td>
<td>3</td>
<td>5.8</td>
<td>5</td>
<td>6.7</td>
<td>7</td>
<td>9.3</td>
<td>10</td>
</tr>
<tr>
<td>June assessment focus</td>
<td>3</td>
<td>5.7</td>
<td>1</td>
<td>1.9</td>
<td>5</td>
<td>9.4</td>
<td>10</td>
</tr>
<tr>
<td>Time allocation</td>
<td>10</td>
<td>19.2</td>
<td>11</td>
<td>21.2</td>
<td>8</td>
<td>15.4</td>
<td>10</td>
</tr>
<tr>
<td>Sequence of learning</td>
<td>1</td>
<td>1.9</td>
<td>2</td>
<td>3.8</td>
<td>2</td>
<td>3.8</td>
<td>8</td>
</tr>
<tr>
<td>Preparing</td>
<td>4</td>
<td>7.4</td>
<td>3</td>
<td>5.6</td>
<td>4</td>
<td>7.4</td>
<td>11</td>
</tr>
<tr>
<td>Implementation advice</td>
<td>2</td>
<td>3.7</td>
<td>1</td>
<td>1.9</td>
<td>3</td>
<td>5.6</td>
<td>11</td>
</tr>
<tr>
<td>Suggested implementation plan – Reading and viewing</td>
<td>1</td>
<td>2.7</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>13.5</td>
<td>5</td>
</tr>
<tr>
<td>Suggested implementation plan – Writing and designing</td>
<td>1</td>
<td>2.6</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>2.6</td>
<td>3</td>
</tr>
<tr>
<td>Making judgments</td>
<td>1</td>
<td>1.9</td>
<td>1</td>
<td>1.9</td>
<td>3</td>
<td>5.8</td>
<td>11</td>
</tr>
<tr>
<td>Using the data for planning</td>
<td>3</td>
<td>5.6</td>
<td>1</td>
<td>1.9</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>Future directions</td>
<td>3</td>
<td>5.7</td>
<td>1</td>
<td>1.9</td>
<td>1</td>
<td>1.9</td>
<td>8</td>
</tr>
<tr>
<td>Appendix A Book talk</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>17.5</td>
<td>2</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>Appendix B Reading running record</td>
<td>4</td>
<td>9.8</td>
<td>3</td>
<td>7.3</td>
<td>2</td>
<td>4.9</td>
<td>2</td>
</tr>
<tr>
<td>Appendix C Comprehension Questions</td>
<td>1</td>
<td>2.4</td>
<td>4</td>
<td>9.8</td>
<td>1</td>
<td>2.4</td>
<td>3</td>
</tr>
<tr>
<td>Appendix D – Work samples</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

To continue this example of the responses collected about the June Checkpoint Assessments, Figure 4.5 below represents the distribution of responses to the Implementation Plan – Reading, and Figure 4.6 represents the distribution of responses to the Implementation Plan – Writing. Out of the participants who implemented The Diverse Family - Reading, 2.7% (n = 1)
found the suggested implementation plan – reading not at all useful. The majority of the participants selected rank numbers 5 and 6, with 29.7% \((n = 11)\) identifying with these scale numbers. Overall, 10.8% \((n = 4)\) of participants found this plan extremely useful.

![Figure 4.5: Distribution of responses across a 7 point Likert scale of the usefulness of the Implementation Plan – Reading within the Diverse Family Assessment represented as percentages](image)

Out of the participants who implemented The Diverse Family – Writing and Designing, 2.6% \((n = 1)\) found the suggested implementation plan – writing and designing not at all useful. The majority of the participants \((44.7%; n = 17)\) selected rank number 5. Overall, 10.5% \((n = 4)\) of participants found this plan extremely useful.
Figure 4.6: Distribution of responses across a 7 point Likert scale of the usefulness of the Implementation Plan – Writing and Designing within the Diverse Family Assessment represented as percentages

It is possible to see demonstrated in these representations that teachers' responses to the usefulness of the implementation plans of these Assessments was relatively positive. The only visible differences to this pattern of relatively positive responses to the June Checkpoints Assessments Components relates to the teachers’ responses to Post Office – Number and Round the Ridges – Giving Directions and Finding Pathways. Table 4.14 below reports the distribution of responses across a 7 point Likert scale (not at all useful to extremely useful) when respondents were asked to comment on the usefulness of each component of the Post Office Number Assessment.
Table 4.14: Distribution of responses (n & %) across a 7 point Likert scale when asked to comment on the usefulness of components of each Assessment in the Post Office - Number June Checkpoint.

<table>
<thead>
<tr>
<th></th>
<th>Not at all useful</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Extremely useful</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>4</td>
<td>13.3</td>
<td>-</td>
<td>6</td>
<td>7</td>
<td>23.3</td>
<td>7</td>
</tr>
<tr>
<td><strong>June assessment focus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>2</td>
<td>7.4</td>
<td>9</td>
<td>33.3</td>
</tr>
<tr>
<td><strong>Time allocation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>3</td>
<td>11.1</td>
<td>5</td>
<td>18.5</td>
<td>4</td>
<td>14.8</td>
<td>7</td>
</tr>
<tr>
<td><strong>Sequence of learning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>2</td>
<td>7.4</td>
<td>9</td>
<td>33.3</td>
</tr>
<tr>
<td><strong>Preparing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>1</td>
<td>3.7</td>
<td>2</td>
<td>7.4</td>
<td>8</td>
</tr>
<tr>
<td><strong>Implementation advice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>2</td>
<td>7.4</td>
<td>9</td>
<td>33.3</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Posting letters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.7</td>
<td>1</td>
<td>3.8</td>
<td>1</td>
<td>3.8</td>
<td>8</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Building envelopes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>3</td>
<td>11.5</td>
<td>-</td>
<td>1</td>
<td>3.8</td>
<td>12</td>
<td>46.2</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Making half</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.7</td>
<td>1</td>
<td>3.8</td>
<td>2</td>
<td>7.7</td>
<td>9</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Postie Pam</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>3</td>
<td>11.5</td>
<td>1</td>
<td>3.8</td>
<td>1</td>
<td>3.8</td>
<td>8</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Basic facts and problem solving</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.7</td>
<td>-</td>
<td>1</td>
<td>3.8</td>
<td>12</td>
<td>46.2</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Creating number stories</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>1</td>
<td>3.7</td>
<td>13</td>
<td>48.1</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Stamps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.7</td>
<td>-</td>
<td>4</td>
<td>15.4</td>
<td>9</td>
<td>34.6</td>
</tr>
<tr>
<td><strong>Suggested implementation plan – Paying for stamps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.7</td>
<td>-</td>
<td>1</td>
<td>3.8</td>
<td>12</td>
<td>46.2</td>
</tr>
<tr>
<td><strong>Making judgments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>2</td>
<td>7.4</td>
<td>9</td>
<td>33.3</td>
</tr>
<tr>
<td><strong>Using the data for planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>3</td>
<td>11.1</td>
<td>6</td>
<td>22.2</td>
</tr>
<tr>
<td><strong>Future directions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>7.4</td>
<td>-</td>
<td>1</td>
<td>3.7</td>
<td>7</td>
<td>25.9</td>
</tr>
<tr>
<td><strong>Appendix A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeracy class checklist – Year 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>3</td>
<td>11.1</td>
<td>-</td>
<td>1</td>
<td>3.7</td>
<td>9</td>
<td>33.3</td>
</tr>
</tbody>
</table>

86
The high frequencies of respondents who ranked the implementation plans related to this Assessment, and to the Round the Ridges Assessment – Giving Directions and Finding Pathways (29.6%), at the midpoint of the scale, suggests that there were some issues that teachers felt needed addressing in these particular tasks. Qualitative comments related to both these assessments suggest that the purpose and focus were more difficult to understand in these Assessments, while the confusing nature of the instructions and the tasks themselves is also mentioned. It seems that these tasks had issues related to validity – teachers were not convinced they were assessing what they said they were assessing.

However looking across the data set, the quantitative data collected of teachers’ reports of the usefulness of the various components of the June Assessments is relatively positive. That is to say that teachers rated the components of the Assessments as being useful as they assessed their students. This is not to say that there are not issues with the current forms of the June Assessments that need addressing. These issues include:

**Issues related to clarifying language used, and explicitly communicating the instructions, purposes, intent and link to relevant Indicators;**

- As mentioned above the Year 1 Checkpoints as a suite are in need of a clear statement, and training and advice for teachers, on the understandings of literacy and numeracy that are being deployed within the assessment tools and processes being advocated. This is more specifically the case in terms of the specificities of literacy and numeracy content included in the June Assessments. As an example, understandings of the sound/symbol relationships evident in the English language, and a justification for the approach taken within the Spelling Search and Letter Sort Assessments are called for. Current understandings of the sound symbol relationship would suggest that teaching young children that each letter has a single equivalent sound
is counter to children learning about the complex relationship between the 26 letters and 40-44 sounds of the English language. The Letter Sort assessment explicitly asks students to name the sound of each letter. The Spelling Search Assessment confuses assessing knowledge of sounds with looking for letters and combinations of letters. With a clearly stated understanding of the spelling, and sound/symbol understandings that are the target of the Assessments (What phonemic and phonic understandings are being valued here?) the Assessments could easily be adjusted to clarify purpose and the intent of the tasks (What are we assessing here?) with the Indicator they purport to provide evidence about.

- Several of the assessments are in need of a clear and precise statement of what they assess and how this relates to the Indicator(s) that they were designed to assess. For example the link between the tasks in the Spelling Search and a student's ability to spell high frequency words seems tenuous, and requires further justification and clear instructions for teachers.

- There is a need to provide teachers with clear statements about what is being assessed (it is not always the Indicator, but often only a component of the Indicator) and advice on what this will mean in terms of being able to make judgments about student achievement of the Indicators. As an example, in the writing task of The Diverse Family the task of writing a factual description does not provide a forum for demonstrating the ability to write and read back statements, questions and commands because questions and commands are not features of this genre. What this means is that no student will demonstrate achievement of this Indicator in its entirety through this task. There is a need for a clear and exact statement for teachers about what this means for their judging and reporting.

**Issues related to increasing technical complexity of the tasks by elaborate assessment sequences and contextualizing activities;**

- The technical complexity of a task can be increased as attempts are made to integrate the assessment in context. This may result in the conflation of skills from one domain to another. In patterns in Songs and Games for example the activity to write or draw pictures to represent the rule would seem likely to conflate the understanding being assessed with a complex task of representation, along with other literacy skills such as writing. There is an opportunity for this task to interrupt the demonstration of the actual skills involved in identifying, creating and extending increasing and decreasing patterns rather than assessing them fairly.

- Overt attempts to contextualise the assessments on the basis of constructivist principles of learning may have the result of increasing
the technical complexity of some of the June Assessments. As an example in Patterns in Songs and Games, the songs themselves are no doubt a useful and engaging way to teach students about patterns, but whether they are the most efficient context to assess this ability in terms of ease and simplicity of assessment is questionable. Similarly in the Book Talk assessment tasks of The Diverse Family the boundaries of what is being assessed are blurred – is the task about students’ understandings of families or about reading?

- The Spelling Search tasks are elaborate combinations of complex activities to assess what seems to be a simple skill related to spelling and using words in written form. Again the activities are likely to be engaging learning activities, but as assessment tools seem overly elaborate, in ways that might well impact on validity of the assessment instruments.

Issues related to the extent to which they are fair and reliable instruments for all students;

- No assessment task can be culture free. However it is important to work toward culture fair assessment. The June Assessments currently have relatively middle-class, mono-cultural patterns of life and experience as their content – families order pizza, letters are delivered through post offices, people go shopping, mommies help children play sport. For students for whom these activities are not familiar, those who live in remote Indigenous communities for example, the use of such content is an issue of content validity. The assessments become tests of knowledge and understandings of middle-class life styles and not assessments of literacy and numeracy. Because no assessment item can be culture free, it is important that suites of alternative assessment tools be made available. At least some of the alternatives should be designed and implemented with a focus on culture fairness for the diverse populations of students evident in Queensland. A particular focus on providing fair and equitable assessment for Indigenous students is called for in a context that claims to be working to improve outcomes of Indigenous students in school.

- There is currently little advice given to teachers about using, amending or adapting the Assessment instruments in inclusive ways to provide fair, valid and reliable evidence of the literacy and numeracy achievement of all students and specifically those with specific language, cognitive, physical, behavioural, social and cultural needs for learning. Clear statements of processes to differentiate assessment for these students are required. For example the Running Record (The Diverse Family) claims to offer the opportunity to collect evidence of the students’ ability to use strategies and cues when reading. In reality this is unlikely if the book is easy, or difficult for the student. Many teachers struggled with the purpose of this task across the full range of
readers in their class. This is not to say that the task need be adjusted necessarily, but its purpose needs to be clarified, and there needs to be an explicit justification for this purpose and clear instructions given to teachers about implementing the task to collect evidence, record performance and make judgments about the achievement evidenced by this task for a diverse cohort of students.

The Year 1 Checkpoints aim to support teachers in monitoring children’s progress toward the Year 1 Literacy and Numeracy Indicators to inform planning for teaching learning and assessment. As mentioned above the instruments themselves focus on what a student can do, and not on what a student cannot do. This leaves any diagnostic potential of the tools to rely on what teachers do with the evidence collected, and how it is used to inform future teaching. As such the materials must involve clear and explicit advice about using the evidence to analyse students’ achievement, and planning for future teaching and learning for individuals, small and large groups. It will be important to ensure that the Future Directions advice provides advice at a level of specificity to enable focus to whole class and individual diagnosis of particular teaching and learning needs. The work samples go to filling this gap somewhat, however teacher capacity in these processes is likely to be variable and consideration of the implications of this for teacher advice and training plans is required.

CONSIDERATIONS

Clear statements of foundational understandings or assumptions of both literacy, numeracy and assessment, and justifications for the same, are required both for the Year 1 Checkpoints generally, but also for relevant components and concepts assessed as part of the Assessments. For example a clear statement on the assumptions being made about how students learn spelling and associated concepts should be used to justify the approach taken to assessing students’ achievement of the indicators related to these concepts in the instruments. The design of these justifications would also provide the opportunity for the purposes and intents of the Assessments to be tested and adjustments to be made if necessary.

Explicit links between the task and the Indicator being assessed need to be provided. Highlighting specific components of indicators that are assessed may reduce confusion for teachers about what is being assessed. Providing explicit advice about recording students achievement in educationally useful ways that support diagnostic approaches to the evidence collected would strengthen the diagnostic power of the Assessments generally.

As adjustments are made to the June Assessments consideration of increased technical complexity and issues related to content validity should be considered.
If the Assessments are to be deployed in a manner that is diagnostically useful for individuals or small groups, consideration of the advice and training for teachers that will be required is necessary.

Highlighting components of the Indicator actually assessed by particular tasks could avoid confusion for teachers on exactly what they are assessing. This is likely to help to explicitly identify the domains being assessed.

The design of alternative tasks that assess the targeted Indicators differently but equivalently would support teachers by enabling them to select a task that is relevant to their contexts and students from a bank of alternative assessments. This design process must include the development and design of culturally relevant materials for Indigenous students.

Provision of reading books, copies of songs (scores and words) and other relevant materials may support schools in implementation of the Assessments.

Consideration of when consistency is absolutely necessary and when relevant data can be collected in different ways. For example, is there a justifiable reason for constraining the style and format of the Running Record used?

Assessments Implemented in the October Assessment Checkpoint
The October Assessments include two versions of what are described as ‘comparable tasks’ that work as integrated units of assessment under the organisers of ‘Shopping’ and ‘Shadows are everywhere’. Each of the integrated assessment units target the same selection of Year 1 Indicators. The unit approach of these instruments seems to fit well with teachers’ preconceived ideas of the place of experience and context and how this relates to effective assessment. However as part of the Supported Trial and due to the truncated timelines of this trial, many teachers reported that they were either teaching these units in addition to their previously planned teaching and learning units, or that they had had to put on hold their plans and shift their program to covering the material for the assessments. Again this issue of time and timing will have had implications for how the teachers involved were able to take up the instruments.

The assessment unit approach to the October assessments reflects the attempt to design assessment tools that are part of the normal everyday teaching and learning activities. There is some justification in this idea, particularly when assessing in the early years. However it also raises issues of the need to ensure that the purposes, domains, limits and uses of the instruments and individual tasks are clearly stated. If this is not achieved there is the risk that the ‘integrating content’ will be blurred with the skills and understandings being assessed. Teachers also need to be provided with clear instructions about the scope of the assessment units as teaching plans.
To provide information about the levels of participation in each of the Instruments, teachers were asked to provide information about which of the assessment units that they selected to trial. Within the cohort of year 1 teachers who replied to the survey, 51% \( (n = 31) \) of the respondents reported implementing Shadows are Everywhere, and 49% \( (n = 30) \) reported implementing Shopping (see Table 4.15).

**Table 4.15: Number and percentage of respondents who implemented each of the October Assessments.**

<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( % )</th>
</tr>
</thead>
<tbody>
<tr>
<td>% implemented Shadows are Everywhere</td>
<td>31</td>
<td>50.8</td>
</tr>
<tr>
<td>% implemented Shopping</td>
<td>30</td>
<td>49.2</td>
</tr>
</tbody>
</table>

Respondents were then asked to rate the level of usefulness of components of the Assessment selection that they had implemented with their students, across a 7 point Likert scale (not at all useful to extremely useful). The responses across both Assessments and across all components were generally positive, the mode response in most cases being 5 or 6 (a positive response). However there was a full range of responses. Table 4.16 demonstrates the distribution of responses across the 7 point Likert scale for *Shadows are Everywhere* and Table 4.17 demonstrates the distribution of responses across the 7 point Likert scale for *Shopping* when teachers were asked to rate the usefulness of the components of the assessment units.
<table>
<thead>
<tr>
<th>Component</th>
<th>Not at all useful</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Extremely Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1: Plan the investigation</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>5 16.1</td>
<td>10 32.3</td>
<td>10 32.3</td>
<td>3 9.7</td>
<td></td>
</tr>
<tr>
<td>Section 2: Dictated spelling</td>
<td>2 6.5</td>
<td>2 6.5</td>
<td>6 19.4</td>
<td>4 12.9</td>
<td>5 16.1</td>
<td>11 35.5</td>
<td>2 6.5</td>
</tr>
<tr>
<td>Section 3: Measure shadows with paper streamers</td>
<td>3 9.7</td>
<td>6 19.4</td>
<td>4 12.9</td>
<td>5 16.1</td>
<td>11 35.5</td>
<td>2 6.5</td>
<td></td>
</tr>
<tr>
<td>Section 4: Estimate length of the shadow streamer</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>2 6.5</td>
<td>6 19.4</td>
<td>8 25.8</td>
<td>10 32.3</td>
<td>2 6.5</td>
</tr>
<tr>
<td>Section 5: Record and report on the observations and compare lengths</td>
<td>3 9.7</td>
<td>2 6.5</td>
<td>3 9.7</td>
<td>4 12.9</td>
<td>7 22.6</td>
<td>10 32.3</td>
<td>2 6.5</td>
</tr>
<tr>
<td>Section 6: Add and take away lengths</td>
<td>3 9.7</td>
<td>2 6.5</td>
<td>6 19.4</td>
<td>7 22.6</td>
<td>6 19.4</td>
<td>5 16.1</td>
<td>2 6.5</td>
</tr>
<tr>
<td>Section 7: Order three lengths</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>6 19.4</td>
<td>8 25.8</td>
<td>11 35.5</td>
<td>3 9.7</td>
<td></td>
</tr>
<tr>
<td>Section 8: Make half</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>5 16.1</td>
<td>8 25.8</td>
<td>12 38.7</td>
</tr>
<tr>
<td>Section 9: Write a factual report</td>
<td>1 3.2</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>5 16.1</td>
<td>7 22.6</td>
<td>13 41.9</td>
</tr>
<tr>
<td>Section 10: Book talk and reading comprehending a factual text</td>
<td>3 9.7</td>
<td>1 3.2</td>
<td>4 12.9</td>
<td>5 16.1</td>
<td>5 16.1</td>
<td>19 32.3</td>
<td>3 9.7</td>
</tr>
<tr>
<td>Appendix A: Dictated spelling</td>
<td>3 9.7</td>
<td>1 3.2</td>
<td>4 12.9</td>
<td>7 22.6</td>
<td>5 16.1</td>
<td>7 22.6</td>
<td>4 12.9</td>
</tr>
<tr>
<td>Appendix B: Book Talk</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>8 25.8</td>
<td>3 9.7</td>
<td>12 38.7</td>
<td>3 9.7</td>
</tr>
<tr>
<td>Appendix C: Reading running record</td>
<td>2 6.5</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>7 22.6</td>
<td>3 9.7</td>
<td>19 32.3</td>
<td>3 9.7</td>
</tr>
<tr>
<td>Appendix D: Comprehension questions</td>
<td>1 3.2</td>
<td>1 3.2</td>
<td>2 6.5</td>
<td>9 29.0</td>
<td>4 12.9</td>
<td>11 35.5</td>
<td>3 9.7</td>
</tr>
<tr>
<td>Appendix E: Work Sample – writing</td>
<td>1 3.2</td>
<td>4 12.9</td>
<td>6 19.4</td>
<td>5 16.1</td>
<td>11 35.5</td>
<td>4 12.9</td>
<td></td>
</tr>
<tr>
<td>Appendix F: Numeracy class checklist</td>
<td>3 9.7</td>
<td>2 7.4</td>
<td>6 19.4</td>
<td>7 22.6</td>
<td>5 16.1</td>
<td>10 32.3</td>
<td>4 12.9</td>
</tr>
<tr>
<td>Additional Assessments – Decorating</td>
<td>2 7.4</td>
<td>2 7.4</td>
<td>9 33.3</td>
<td>6 22.2</td>
<td>6 22.2</td>
<td>6 22.2</td>
<td>2 7.4</td>
</tr>
<tr>
<td>Additional Assessments – What can I buy?</td>
<td>4 15.4</td>
<td>2 7.7</td>
<td>1 3.8</td>
<td>6 23.1</td>
<td>6 23.1</td>
<td>5 19.2</td>
<td>2 7.7</td>
</tr>
<tr>
<td>Additional Assessments – Where are you?</td>
<td>4 14.3</td>
<td>1 3.6</td>
<td>2 7.1</td>
<td>8 28.6</td>
<td>5 17.9</td>
<td>6 21.4</td>
<td>2 7.1</td>
</tr>
<tr>
<td>Additional Assessments – What time is it?</td>
<td>3 10.3</td>
<td>1 3.4</td>
<td>3 10.3</td>
<td>8 27.6</td>
<td>5 17.2</td>
<td>7 24.1</td>
<td>2 6.9</td>
</tr>
</tbody>
</table>
Table 4.17: Distribution of responses across the 7 point Likert scale for Shopping when asked to comment on the usefulness of each component. Shading represents the most popular code selected.

| Section | Not at all useful | | 2 | 3 | 4 | 5 | 6 | Extremely Useful |
|---------|------------------|---|---|---|---|---|------------------|
| Maisy Goes Shopping (Cousins, 2001) | 4 | 13.3 | 2 | 6.7 | 2 | 6.7 | 3 | 10.0 | 5 | 16.7 | 11 | 36.7 | 3 | 10.0 |
| Section 1. Dictated spelling – shopping list | 1 | 3.3 | - | - | 2 | 6.7 | 4 | 13.3 | 4 | 13.3 | 12 | 40.0 | 7 | 23.3 |
| Section 2. Book talk, reading and comprehending | 1 | 3.3 | 2 | 6.7 | 1 | 3.3 | 3 | 10.0 | 6 | 20.0 | 12 | 40.0 | 5 | 16.7 |
| Section 3. Writing a recount of a shopping experience | - | - | - | - | - | - | - | 5 | 17.2 | 10 | 34.5 | 9 | 31.0 | 5 | 17.2 |
| Section 4. Fruit shop – count and record numbers of fruit using diagrams | 1 | 3.3 | 3 | 10.0 | 1 | 3.3 | 5 | 16.7 | 5 | 16.7 | 11 | 36.7 | 4 | 13.3 |
| Section 5. Identify and compare whole numbers | - | - | - | - | 3 | 10.0 | 5 | 6.7 | 7 | 23.3 | 13 | 43.3 | 2 | 2.7 |
| Section 6. Order and position | - | - | - | - | 1 | 3.4 | 5 | 17.2 | 9 | 31.0 | 9 | 31.0 | 5 | 17.2 |
| Section 7. Solve addition and take away problems/Create addition and take away stories | - | - | - | - | 1 | 3.4 | 4 | 13.8 | 9 | 31.0 | 11 | 37.9 | 4 | 13.8 |
| Section 8. Maintain equivalence – Amounts to $1 using coins | - | - | - | - | 1 | 3.3 | 4 | 13.3 | 6 | 20.0 | 14 | 46.7 | 5 | 16.7 |
| Section 9. Demonstrate half a collection | - | - | - | - | 1 | 3.4 | 4 | 13.8 | 4 | 13.8 | 16 | 55.2 | 4 | 13.8 |
| Section 10. Direct comparisons – Distinguishing the differences in mass by hefting | - | - | - | - | - | - | 6 | 20.0 | 5 | 16.7 | 15 | 50.0 | 4 | 13.3 |
| Section 11. Order mass – Distinguishing the differences in mass by hefting | - | - | 1 | 3.3 | 2 | 6.7 | 6 | 20.0 | 6 | 20.0 | 12 | 40.0 | 3 | 10.0 |
| Section 12. Make mind pictures | 1 | 3.4 | 1 | 3.4 | 1 | 3.4 | 4 | 13.8 | 8 | 27.6 | 10 | 34.5 | 4 | 13.8 |
Again, while there are a full range of responses to the request to rate the usefulness of the instrument components it is also evident from the tables above that there was generally a positive response from the Year 1 teachers who responded. The qualitative comments collected produced data similar to that collected on the June Assessments. In addition teachers particularly noted issues with the bias toward individual one-to-one assessment techniques, and the quantity of assessment required to be completed. There were also issues raised about the appropriateness of the content organisers of the instruments in relation to cultural fairness, practical difficulties with finding light for shadows, and resource issues. Time and timing featured as key concerns as has been the case with the other Checkpoints.

Teachers were also concerned with issues related to the instruments and their use with students who they felt would easily achieve the targeted Indicators, or those who they believed would struggle. This seems to call on a need to clarify the purpose of the checkpoint as a way to validate teacher judgments rather than to identify students for future diagnostic intervention.

The October Assessments are detailed and seem to fit well within the boundaries of early years’ classroom pedagogy. This has no doubt encouraged teachers to see them as useful, and without the late notification of the need to complete a ‘unit’ of work would be likely to be utilised within normal classroom units without a great deal of difficulty. Teachers raised issues of the need for alternatives in contexts where multi-age classes require three and four year cycles of units and content coverage. The inclusion of at least one
additional, and possibly more, October Assessment tools would also allow for issues related to culture fairness, and the needs of specific contexts to be addressed. The choice of just two units caused issues for many participants. For while the content of 'Shopping' was not necessarily relevant to students in remote communities in the north of Queensland (some communities do not have shops and so shopping is experienced differently), selecting the Shadows Assessment brought issues related to climate to the fore (there are long periods when Shadows are not easily captured during the wet season, or in the Summer months). While teachers found ways around this in the Supported Trial, the provision of other equivalent tasks would address such issues with greater consistency.

The amount of information about students' achievement of the expectations detailed in the year 1 indicators through either of these Assessment units is expansive. There are useful supports and samples to detail ways to organise this information as part of the materials. The process of moderation also supported these processes in the 2010 Supported trial.

Teachers expressed concern with the amount of time required for Assessment in each of the October Assessments units. While assessment is recognised as an important part of teaching and learning, it is likely that there are alternatives to individual one-to-one assessment that could be used to collect evidence about students' achievement of some of the targeted Indicators at least. While the focus on everyday literacy and numeracy tasks is valid, the expectation of so much teacher time being spent conducting one-to-one assessment may not be realistic – or perhaps educationally sound - in busy classrooms.

Issues related to the format of Running Records and the naming of cues and strategies were of concern to teachers also. We suggest that those tasked with continued development of the Assessment tools consider the requirement of consistency and how this balances with teacher efficiency of assessing the literacy achievements of students. Is it important that all teachers use the same Running Record format? Of more concern to our analysis of the reading and viewing indicators are issues related to book selection. Having all students read the same text does provide particular understandings about student achievement and may be justifiable. However we recommend again that if the Assessments continue to be structured in this way that there needs to be a justification for this made explicitly to teachers. Taking such an approach does limit the capacity of teachers to collect information about students' use of strategies if a child either finds the text easy or difficult to read. Teachers expressed concern about this, and there are some sound reasons for assessing students reading a text within their instructional reading capabilities. In the end the two tasks –reading a text within your instructional level, and reading the same text as your classmates regardless of the book's level of difficulty to you as a reader - achieve different purposes. However there is enough evidence in the data to suggest that the purposes of
the approach taken are not clear to teachers, and this needs to be addressed through communication of a clear justification.

Additionally but still in relation to text selection for the running record is the use of the Maisey goes Shopping (Cousins, 2001) text in the 'Shopping' Assessment. This text is a mixed-genre or hybrid text that combines a complex variety of interactional sequences, grammatical mood shifts, and implied interactional directions. The use of this text would seem to be about content and perhaps the inclusion of a 'real world' text. The implications of this overt link to progressive notions of literacy are that students are called upon to read a text that has a difficulty level well beyond that suggested by the format, style and vocabulary inclusions. The text type also does not fit neatly as an expert model for the writing task that comes after reading. There were similar issues with the Mummy Book (Parr, 2002) in the June Assessments.

Teachers expressed real concern about the link between WD1xiii indicator with the Spelling task and lists of acceptable spellings (for both Shadows are Everywhere and Shopping). Teachers concerns were about what they considered to be seemingly arbitrary selections of some misspellings as acceptable and the exclusion of others. While the acceptable spellings have attempted to focus on sound symbol relationships, there are still acceptable alternatives that are excluded (lion for line as an example). The task could be adjusted to provide teachers with explicit justification and principles for deciding the acceptability of alternatives, rather than a list of alternatives. Additionally by providing advice and training on how to use the analysis of misspellings to collect relevant diagnostic information, rather than artificially constraining the acceptable alternatives, the task would support classroom practice in useful ways.

CONSIDERATIONS

Consider issues related to text selection and the use of the same text to contextualise the activity and as the reading task.

Consider the balance between consistency and allowing teachers to draw on their skills and expertise. This is particularly pertinent with regard to the running record formats used. Perhaps providing alternatives, along with training as support for those not experienced in conducting running records would be more useful than mandating a particular format.

Consider the spelling task and its link to the Indicator that it is fashioned to assess. Are there more useful and efficient ways to assess students' achievement of this indicator?

October Checkpoints – As a Diagnostic Tool

From the Year 1 teacher sample of participants, 14.8% \((n = 9)\) Strongly agreed that the October Checkpoints would be useful as a diagnostic tool. However, 6.6% \((n = 4)\) of the sample strongly disagreed with this statement.
This is represented in Table 4.18 below. From this same sample, 13.1% \((n = 8)\) *Strongly agreed* that the October Checkpoints would be useful as an assessment process. However, 8.2% \((n = 5)\) of the sample *strongly disagreed* with this statement. The most selected response in both cases was 6, suggesting a generally positive response.

**Table 4.18: Distribution of responses across a 7 point Likert scale when Year 1 teachers were asked to comment on their level of agreement (strongly disagree to strongly agree) on the use of the October Assessments as a diagnostic tool, or more generally as an early years assessment process.**

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
</tr>
<tr>
<td>Useful – Diagnostic Tool</td>
<td>4</td>
<td>6.6</td>
<td>3</td>
<td>4.9</td>
<td>5</td>
<td>8.2</td>
<td>7</td>
</tr>
<tr>
<td>Useful – Assessment Process</td>
<td>5</td>
<td>8.2</td>
<td>3</td>
<td>4.9</td>
<td>5</td>
<td>8.2</td>
<td>7</td>
</tr>
</tbody>
</table>

From the Facilitators sample of participants, 27.3% \((n = 6)\) *Strongly agreed* that the October Checkpoints would be useful as a diagnostic tool, and no respondents suggested that they strongly disagreed with this statement. This is represented in Table 4.19 below. From this same sample, 13.6% \((n = 3)\) *Strongly agreed* that the October Checkpoints would be useful as an assessment process, and again none of the sample *strongly disagreed* with this statement. The most likely to be selected response in both cases was 6 (and 5 in the case of useful as a diagnostic tool), suggesting a generally positive response.

**Table 4.19: Distribution of responses across a 7 point Likert scale when Facilitators were asked to comment on their level of agreement (strongly disagree to strongly agree) on the use of the October Assessments as a diagnostic tool, or more generally as an early years assessment process.**

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
</tr>
<tr>
<td>Useful – Diagnostic Tool</td>
<td>2</td>
<td>9.1</td>
<td>7</td>
<td>31.8</td>
<td>7</td>
<td>31.8</td>
<td>6</td>
</tr>
<tr>
<td>Useful – Assessment Process</td>
<td>1</td>
<td>4.5</td>
<td>2</td>
<td>9.1</td>
<td>8</td>
<td>36.4</td>
<td>8</td>
</tr>
</tbody>
</table>
Conclusion

The proceeding analysis of the Year 1 Checkpoints Assessments was conducted in three stages:

- The review of the instruments\(^5\);
- The use of surveys to collect information from participants in the Supported Trial, and the descriptive quantitative analysis of this data for trends; and
- The coding of qualitative data collected via surveys and interviews, and work samples collected from participants to exemplify the trends identified.

The findings presented here are exemplary only, limited by the duration of the project, timing issues that had implications for the sample size of respondents to surveys, and the consequent quantitative analysis possible. The excess of qualitative data collected worked in some ways to address the small survey sample sizes but many of the qualitative responses addressed very specific technical issues with particular sections of the Assessments. This data will be passed to those redrafting and developing the Year 1 Checkpoints, but is limited in its ability to support a review of the materials at the level required here. We offer the following findings and recommendations from the analysis that has been possible:

- That the decision to trial the materials in a truncated timeline had implications for both the teachers involved in the Supported Trial and for the Evaluation of the Supported Trial.
- That the Indicators require review in light of their domain specificity, content validity and issues related to their ability to record students’ achievement of end of year expectations in fair, valid and reliable ways.
- That the Indicators are framed with a focus on end of year expectations and this provides the possibility that this framework can work to focus the assessment process on informing teaching rather than describing the individual capacities of children against normative descriptors. The Indicators can be used to discuss what students have and have not been taught, what they have and have not learnt, rather than map their progress on a singular developmental pathway. The distinction suggested here is fine, and easily lost, especially as the actual differences between these Indicators and a developmental continuum such as those currently used in the Year 2 Diagnostic Net are much more related to framing than content or form. The Indicators do not in

---

\(^5\) This review was supported by the expert comments of Karen Dooley (QUT) and Beryl Exley (QUT). Their input to the analysis of the materials presented in this chapter is acknowledged here.
their current form lend themselves to being used to identify students for interventions or further resourcing.

- The links between the Checkpoints, the targeted Indicators, the Indicators more generally and assumptions about literacy, numeracy, language and learning are not explicitly stated in the materials. As the suite is further developed and if deployed as an early years assessment tool this should be addressed.

- There is no advice given to teachers about how the information collected should be reported to students and parents, and this should be a focus of future materials.

- Many teachers generally saw the Year 1 Checkpoints as useful classroom based assessment items or tasks. Teachers expressed issues with the suite and these have been clearly laid out in this chapter.

- That the Year 1 Assessment Checkpoints require front loading materials to ensure that: their approach to literacy and numeracy is explicitly stated; advice and training on differentiating assessment for the specific needs of students be provided; advice and training on issues related to culture fair assessment be provided; and that advice and training on the use of evidence collected to make judgments about student achievement and to analyse and use evidence to inform future teaching and learning be provided.

- The February Assessments are currently lists of effective advice about everyday classroom contexts and purposes that can be used as quality assessment opportunities. As such they are adequate and solid. Without appropriate adjunct support materials included their validity and reliability is solely reliant on the professional expertise and capacity of teachers. Without advice provided about future directions or the use of the information collected especially for those students who are not able to achieve the targeted Indicators their use to inform future teaching is not direct.

- The June Assessments have issues related to clarity of their purpose, foundations in assessment and literacy and numeracy, content validity, unnecessary technical complexity as a result of overt attempts to contextualise and embed the assessments in integrated classroom activities, and a lack of clarity in some tasks about their purposes and intents. Specific issues for the redrafting of the materials include a re-assessment of:
  - The Spelling Search and Letter Sort Assessments and their links to current understandings of sound/symbol relations and how children learn spelling, letter and sound skills and understandings;
  - The elaborate procedure of the activity in Spelling Search;
(and a justification provided to teachers for) having all students read the one text and how this relates to collecting evidence to make judgments about achievement of the relevant indicators;

- The Running Record format used, and a decision made about whether consistency is required across all teachers;

- The complexity of context and clarity of instructions in Patterns in Songs and Games;

- The technical complexity created across the suite of assessments by embedding the tasks in sometimes overly elaborate or contextualized activities;

- The Future Directions to consider the level of specificity required to support teachers’ diagnostic deployment of the instrument; and

- The time allocation required to complete all assessments in the current formats.

- The October Assessments as units of assessment seemed to fit within teachers’ understandings of early years assessment and pedagogy and as such were reported by many to be useful as assessment processes and diagnostic tools by the teachers involved in the Supported Trial. They are presented as summative checks of students’ achievement of the targeted Year 1 Indicators. There are no details of future directions for teaching which limits their ability to act formatively. Specific issues for the redrafting of the materials include a re-assessment of:

  - Relevant issues from feedback provided on the June Assessments;
  - Text selection; and
  - The activity designed to assess WD1xiii

The eventual aim should be to provide teachers with a bank of assessments from which to select instruments relevant to their students, contexts and resources, and as such we recommend that alternative equivalent assessments be drafted.

While the analysis in this chapter has uncovered several issues that require attention in relation to the Year 1 Checkpoints Assessments as instruments, and has provided feed forward advice about the framework used in this assessment package – the P-3 indicators – the analysis should not be taken to detail an approach to assessment that is flawed. The focus on achievement of end of year level expectations, the attempt to design assessments that engage students in everyday teaching and learning activities, and the embedding of these assessments in effective early years pedagogy and assessment practices means that the Year 1 Checkpoints, with continued development and refinement have the potential to provide an effective assessment tool for use in
Year 1. To achieve this status issues of workload for teachers and students also need consideration.

This report continues to argue that the ultimate aim of those involved in this process should be to design an effective systems-based approach to assessment across the P-3 years. This systems-based approach should work to wrap effective assessment for a variety of purposes around the National Assessment initiatives mandated for Year 3. The approach as a whole should aim to build assessment capacity in the early years, provide fair, valid and reliable information about students’ achievement to relevant stakeholders, and inform teaching and learning for all students.
5. Moderation as capacity building

**Introduction**

Effectively bridging the purposes and uses of system level accountability initiatives and teacher assessment practices is reliant on teachers’ capacity to interpret assessment data and make consistent and valid judgments about students’ work. Wyatt-Smith, Klenowski & Gunn (2010) remind us that while the use of standards to make teacher judgments about student performance and processes of moderation have been “at the heart of the Queensland senior assessment model” (p. 59) for some time, the use of standards and moderation at other year levels in Queensland has been less central to assessment and teachers’ work. Recently the Queensland Comparable Assessment Task (QCAT) processes, and over a longer time period the moderation processes of the Year 2 Diagnostic Net are examples where teachers in early and middle years of schooling have been expected to be involved in moderation processes. The Year 1 Checkpoints also require teachers to moderate their judgments on the October suite of assessment tasks. This marks one of the first times that Year 1 teachers have been required to undertake interschool moderation as part of a systemic initiative in Queensland. The 2010 trial of these materials saw the first iteration of these processes with the teachers involved in the Supported Trial.

The processes that teachers engage in to moderate judgments made on students’ work and standards are generally under researched. While there are considerable ‘how to’ accounts across policy, research and handbooks, there has been to date little engagement with the actual enactment by teachers of standards and consistent judgments within moderation processes. Some research related to senior schooling (see for example Murphy, 2004) is available, and more recently the work of Klenowski and Wyatt-Smith (see for example Klenowski and Wyatt-Smith, 2010) has focused on middle years’ teachers and moderation processes related to QCATs in Queensland. However the dearth of research into this issue within the field of early childhood, or the early years of schooling is of concern.

Effective assessment practices can support improved student learning and outcomes. However assessment is a complex social practice and involves teacher judgment. There is a challenge to ensure that teacher judgments made are fair and equitable, valid and consistent. Teacher moderation has been identified as an influential strategy for developing the link between assessment and improved practice as well as promoting and supporting consistency of teacher judgment. Moderation is also recognised as effective in building teacher capacity in assessment (Klenowski and Wyatt-Smith, 2010). The process of moderation often involves teachers in shared discussions of
student work, usually in relation to predetermined assessment criteria or standards.

In this way moderation is the process whereby teachers share evidence of their expectations for student work and their understandings of standards, learning and assessment items and processes with other teachers. The purpose is to achieve consistent teacher judgments and improve the consistency of their decisions about student learning and achievement. The moderation process helps teachers make dependable decisions about student progress and achievement, because they are required to share their interpretations with other teachers. When the process involves groups of teachers meeting to share previously assessed student work samples, to negotiate the grade awarded or decision made about that student work as a way to reach consensus, the process is referred to as social moderation (Linn, 1993 as cited in Wyatt-Smith, et al, 2010). This process works to develop a common understanding of the standards, expectations, or milestone descriptors being used, and focuses on similar or consistent recognition of work that reaches those standards, expectations or milestones (Wyatt-Smith, Klenowski & Gunn, 2010). Moderation as judgment practice is fundamental to quality reporting of student achievement. When teachers are given the opportunity to use their own judgments of assessment data with those of other teachers, they are more able to produce valid and reliable judgments that are consistent with one another and with stated standards of performance or statements of achievement (Wyatt-Smith Klenowski & Gunn, 2010).

“Teachers who engage consistently in moderation processes are better able to:

- Assess student performance consistently, effectively, confidently, and fairly;
- Build common knowledge about curriculum expectation and levels of achievement;
- Identify strengths and areas for growth based on evidence of student learning;
- Adjust and acquire new learning by comparing one’s thinking to that of another teacher;
- Share effective practices to meet the need of all students, monitor progress and celebrate growth.”

(The Ontario Literacy and Numeracy Secretariat Capacity Building Series, 2007 p. 1, drawing on the work of Little et al (2003)).

The Ontario Literacy and Numeracy Secretariat Capacity Building Series (2007, p. 5-7) cites four distinct benefits of moderation that align with Little et al (2003).
**Consistency and reliability:** by confirming and validating student performance through moderation processes teachers develop a common understanding of the criteria that are being applied and the levels of achievement outlined by the curriculum documents.

**Collaborative planning:** working together to develop consistent standards for students and consistent judgment of student work encourages a deep understanding of the expectation, standard or criteria being used.

**Fairness and Equity:** building a solid understanding of what is expected from students broadly, provides teachers with understandings necessary for working from an other than deficit frame.

**Alignment of instruction:** de-privatising teacher pedagogic and assessment practice during moderation provides an opportunity for teachers to share strategies. Moderation fosters the alignment of instruction as teachers share strategies among team members. Students are consistently surrounded by effective strategies, based on regular teacher moderation inquiry.

In a context of accountability, teacher assessment will only be considered credible if there are mechanisms in place that support consistent judgments across teachers and schools. In a report from the UK’s Qualifications and Curriculum Authority (Qualifications and Curriculum Authority, 2009) that aimed to investigate accountability processes, the authors made the claim that “taking part in real moderation proved to be a highly effective way of learning how to improve the quality of teachers’ own assessment as well as how to confirm the assessments of others. All the available evidence showed a positive impact on the nature and range of evidence used to support assessment, the accuracy of assessments and the understanding of what characterises performance at a national curriculum level.” (Qualifications and Curriculum Authority 2009, pp.3, 24). Klenowski and Wyatt-Smith support these findings:

> Moderation too is intrinsic to efforts by the profession to realise judgments that are defensible, dependable and open to scrutiny. Moderation can no longer be considered an optional extra and requires system-level support, especially if, as intended, the standards are linked to system-wide efforts to improve student learning.”

(Klenowski & Wyatt-Smith, 2008, p. 1)

However, like any initiatives that require change on the part of teachers, a process of social moderation will only be effective if the system provides appropriate supports (Queensland Teachers Union, 2009) and considers relevant delivery standards that according to the QSA include the provision of:

- Syllabuses or curriculum documents that clearly describe content and achievement standards;
- Contextualized exemplar assessment instruments;
• Samples of student work annotated to explain how they represent different standards of performance;
• Consensus through teacher discussions (social moderation processes) on the quality of the assessment instruments and the standards of student work;
• Professional development of teachers;
• An organisational infrastructure to ensure the above takes place.

(Queensland Studies Authority 2009, p.3)

Based on this review (and particularly The Ontario Literacy and Numeracy Secretariat Capacity Building Series, 2007), the following evaluation framework is called upon to evaluate the effectiveness of moderation processes.

**Consistency and reliability** – moderation provides opportunities for teachers to discuss the standards or expectations being used as the basis for teacher judgment of student work. It is likely that this will occur during discussions of student work samples. Additionally there should be an expectation that teachers will take time to discuss the actual standards, milestones or expectations that are being applied. This means that teachers should as a result of moderation processes in the Year 1 Checkpoints develop a deeper understanding of the Year 1 Indicators and their positioning in the P-3(9) indicators more generally.

**Collaboration and shared planning** – moderation should provide opportunities for teachers to develop shared understandings about not just the expectations (Indicators), but also sound processes for identification of specific student needs and strategies for interventions and future teaching based on the results of assessment. This builds teacher capacity in effective teaching and learning strategies and supports the alignment of approaches across teachers and sites.

**Clear and simple standards or expectations** – moderation is reliant on the foundation that teachers will work together to develop clear and simple understandings of standards and expectations. However the format should also provide clear instructions and resources that unpack the standards and expectations held for teachers as participants in the moderation.

**High equity approaches** – moderation allows for access to a broad range of student performance and achievement standards. Teachers see their own students’ work within the broader scope of students’ work across sites. Moderation processes must work to support teachers in building a solid understanding of what can be expected from their students, and when consistently linked to discussions and understandings of teaching and assessing this will help to move beyond deficit explanations of student cohorts.
The Year 1 Checkpoint approach to moderation

The approach to moderation trialled within the Supported Trial provided Year 1 teachers and other support personnel with an opportunity for social moderation. The materials provided, including the order of presentation for the day, enabled a professional development possibility for the participants focused on assessment capacity building as much as on moderation of work samples. This was particularly effective during this trial year when there were no system reporting requirements. It will be important to consider the professional development opportunities afforded at these sessions if the focus of the moderation day shifts in the future.

Members of the Evaluation team attended several of the Moderation Day sessions, and observations were conducted. These days also afforded many opportunities to speak with teachers and administrators about the Year 1 Checkpoints materials and processes more generally. Work samples of students and teachers’ work in the October Assessment Checkpoints were also collected. Participants were surveyed as a way to collect their opinions about the moderation process broadly.

There were twenty Moderation sites for the Year 1 Checkpoints Assessment moderation day in the 2010 Supported trial. One hundred and sixty-eight participants (n=168) across these sites responded to the Moderation Survey. There was a relatively evenly distributed response rate from each Moderation site, especially when actual attendance at each site is taken into account.

Table 5.1 provides the response rate for each Moderation site and an indication of the percentage of the complete sample that these site-based response rates represent.
Table 5.1: Distribution of Participants who responded to Moderation Survey across Moderation Sites

<table>
<thead>
<tr>
<th>Moderation Site Number</th>
<th>n</th>
<th>%</th>
<th>Moderation Site Number</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>16</td>
<td>9.5</td>
<td>S11</td>
<td>8</td>
<td>4.8</td>
</tr>
<tr>
<td>S2</td>
<td>16</td>
<td>9.5</td>
<td>S12</td>
<td>8</td>
<td>4.8</td>
</tr>
<tr>
<td>S3</td>
<td>14</td>
<td>8.3</td>
<td>S13</td>
<td>8</td>
<td>4.8</td>
</tr>
<tr>
<td>S4</td>
<td>12</td>
<td>7.1</td>
<td>S14</td>
<td>6</td>
<td>3.6</td>
</tr>
<tr>
<td>S5</td>
<td>11</td>
<td>6.5</td>
<td>S15</td>
<td>6</td>
<td>3.6</td>
</tr>
<tr>
<td>S6</td>
<td>11</td>
<td>6.5</td>
<td>S16</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>S7</td>
<td>10</td>
<td>6</td>
<td>S17</td>
<td>4</td>
<td>2.4</td>
</tr>
<tr>
<td>S8</td>
<td>9</td>
<td>5.4</td>
<td>S18</td>
<td>4</td>
<td>2.4</td>
</tr>
<tr>
<td>S9</td>
<td>9</td>
<td>5.4</td>
<td>S19</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>S10</td>
<td>8</td>
<td>4.8</td>
<td>S20</td>
<td>1</td>
<td>.6</td>
</tr>
</tbody>
</table>

Participants in the Year 1 Checkpoint Assessment moderation day who responded to the survey were employed in a range of duties. Teachers of sole Year 1 classes composed 67.9% of the sample (n = 114). Teachers of multi-age classes that included a Year 1 Cohort formed 17.86% of the sample (n = 30). Heads of Curriculum accounted for 2.4% (n = 4) of the sample and the remaining participants held varied duties including Deputy Principal/Principal roles (2.4%), Prep Teachers (4.2%) and Key Teachers and/or Learning Support Teachers: Literacy and Numeracy.
Results suggested no significant correlations between responses on the feedback form and site of moderation ($p > .05$). It is important to note the small sample size which may have had an impact upon these results, however analysis of the qualitative comments across the full data set would also suggest that there were no particular issues expressed about particular sites. The role of the support materials provided to facilitators in creating similar experiences across moderation sites is likely to have been relevant in this finding.

**Participants’ Opinions on Moderation Approach**

The survey asked participants to rank their level of agreement across a 7-point Likert scale on nine statements about the Moderation session processes and resources used on the day. Frequencies were analysed to reveal the distribution of responses. Table 5.2 shows the distribution of respondents across the 7-point Likert scale for each statement.
Table 5.2: Distribution of Responses across the 7-point Likert scale on level of agreement with statements about the approach to moderation.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>The moderation day supported my work in assessments of Year 1 students.</td>
<td>2</td>
<td>1.2</td>
<td>6</td>
<td>3.6</td>
<td>8</td>
<td>4.8</td>
<td>27</td>
</tr>
<tr>
<td>The approach to moderation taken during the session was effective for this purpose.</td>
<td>1</td>
<td>0.6</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>3.6</td>
<td>18</td>
</tr>
<tr>
<td>The process of facilitation was effective.</td>
<td>1</td>
<td>.6</td>
<td>2</td>
<td>1.2</td>
<td>7</td>
<td>4.2</td>
<td>12</td>
</tr>
<tr>
<td>The examples in the workbook (e.g. Sample A) helped to support my learning or the learning of other moderation group members about making consistent judgments</td>
<td>4</td>
<td>2.4</td>
<td>9</td>
<td>5.4</td>
<td>9</td>
<td>5.4</td>
<td>34</td>
</tr>
<tr>
<td>The workbook activities (e.g. Sample B, C, D) helped to support my learning or the learning of other moderation members about making consistent judgments</td>
<td>2</td>
<td>1.2</td>
<td>7</td>
<td>4.2</td>
<td>11</td>
<td>6.5</td>
<td>31</td>
</tr>
<tr>
<td>The think, pair, share approach to arriving at a consistent judgment was effective</td>
<td>2</td>
<td>1.2</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>3.6</td>
<td>16</td>
</tr>
<tr>
<td>The example analysis table and DAAR helped to support my learning or the learning of other moderation group members about acting on the evidence and future directions</td>
<td>1</td>
<td>.6</td>
<td>5</td>
<td>3</td>
<td>13</td>
<td>7.7</td>
<td>34</td>
</tr>
<tr>
<td>The activities relating to the DAAR and acting on the evidence helped to support my learning or the learning of other moderation group members.</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>1.2</td>
<td>10</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td>The moderation day was generally useful</td>
<td>3</td>
<td>1.8</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>4.2</td>
<td>14</td>
</tr>
</tbody>
</table>

The following section will introduce these statements from the moderation survey and briefly discuss the distribution of responses.
Statement: *The Moderation Day supported my work in assessment of Year 1 students*

Overall, 1.2% (n = 2) of the participants identified with *Strongly Disagree* for this statement. Both participants attended the one moderation site. Of the participants who responded to this question, 16.9% (n = 45) selected *Strongly Agree* for this statement. The highest number of participants selected ‘6’ on the Likert Scale (27.1%). Figure 5.1 reports the overall distribution of participants’ responses across the 7 point Likert Scale for this statement.

![Distribution of responses in % on the 7-point Likert Scale](chart.png)

**Figure 5.1: Distribution of responses in % on the 7-point Likert Scale**

*The Moderation Day supported my work in assessment of Year 1 students*

Statement: *The approach to moderation taken during the session was effective for this purpose*

One participant, 0.6% of the sample, identified with *Strongly Disagree* for this statement. There were 38 participants (22.6%) who selected *Strongly Agree* for this statement. Forty-nine participants (29.2%) selected “6”, making this the most identified with interval. For the overall distribution of participants, see Figure 5.2.
Figure 5.2: Distribution of responses on the 7-point Likert Scale

*The approach to moderation taken during the session was effective for this purpose*

**Statement: The process of facilitation was effective**

Out of sample, 0.6% (n = 1) of participants selected *Strongly Disagree* for this statement. Interval “6” was the most selected, with 31.1% (n = 52) of participants identifying with it. *Strongly Agree* was selected by 24.6% (n = 41) of the participants who responded to this question. For further information about the distribution of responses to this question on the 7-point Likert scale, see Figure 5.3.
Figure 5.3: Distribution of responses on the 7-point Likert Scale

The process of facilitation was effective

Statement: *The examples in the workbook helped to support my learning or the learning of other moderation group members about making consistent judgments*

Overall, 2.4% of participants ($n = 4$) selected *Strongly Disagree* for this statement. These participants were situated across three Moderation sites. With 14.9% of participants ($n = 25$) selecting *Strongly Agree* for this statement situated across nine Moderation sites, it is unlikely that this relates to any particular approach to using these samples at particular Moderation sites. Interval scale “6” was the most selected number, with 28.6% ($n = 48$) of the sample identifying with this interval. Figure 5.4 provides further information about the distribution of responses on the 7-point Likert scale for this question.
The examples in the workbook helped to support my learning or the learning of other moderation group members about making consistent judgments

Statement: The workbook activities helped to support my learning or the learning of other moderation group members about making consistent judgments

Overall, 1.2% (n=2) of the sample selected *Strongly Disagree* for this question. These participants were from two separate Moderation sites. Almost 12% (11.9%) of the population selected *Strongly Agree* for this statement (n = 20). The most selected rank was ‘6’ (31%, n = 52). For a representation of this distribution of responses, see Figure 5.5 below.
Statement: *The think, pair, share approach to arriving at a consistent judgment was effective*

Two participants (1.2%) identified with *Strongly Disagree* for this statement. These participants were from two separate Moderation sites. Overall, 19% (\(n = 32\)) of participants *Strongly Agreed* with this statement. These participants were from 14 different Moderation sites. The most selected rank was ‘6’ (36.3%; \(n = 61\)). Figure 5.6 presents a diagrammatic representation of the distribution of responses on the 7-point Likert scale for this statement.
Figure 5.6: Distribution of responses on the 7-point Likert Scale

*The think, pair, share approach to arriving at a consistent judgment was effective*

**Statement:** *The example analysis table and DAAR helped to support my learning or the learning of other moderation group members about acting on the evidence and future directions*

Overall, 0.6% \((n = 1)\) of the sample selected *Strongly Disagree*. There were 21 participants (12.5%) that selected *Strongly Agree*. The most selected rank was ‘6’ (28.6%; \(n = 48\)). For further information about the distribution of responses for this statement, see Figure 5.7 below.
The example analysis table and DAAR helped to support my learning or the learning of other moderation group members about acting on the evidence and future directions.

**Statement:** The activities related to the DAAR and acting on the evidence helped to support my learning or the learning of other moderation group members.

Overall no participants identified with *Strongly Disagree* for this statement. Out of all participants, 11.9% (n = 20) identified with *Strongly Agree*. The most selected rank was ‘6’ (30.5%; n = 51). Figure 5.8 gives further information about the distribution of responses on the 7-point Likert scale for this statement.
The activities related to the DAAR and acting on the evidence helped to support my learning or the learning of other moderation group members

Percentage of Participants

Statement: The moderation day was generally useful

Out of the sample, 1.8% (n = 3) Strongly Disagreed with this statement. These participants were from three separate Moderation sites. Twenty-four percent (n = 40) identified with Strongly Agree for this statement. The most selected rank was ‘6’ (32.9%; n = 55). For further information on the distribution of responses on the 7-point Likert scale, see Figure 5.9 below.

Figure 5.8: Distribution of responses on the 7-point Likert Scale

The activities related to the DAAR and acting on the evidence helped to support my learning or the learning of other moderation group members
Figure 5.9: Distribution of responses on the 7-point Likert Scale

The moderation day was generally useful

These results suggest that general perceptions about the appropriateness, effectiveness and utility of the Moderation day activities were positive as represented by those teachers, leaders and others who responded to the Moderation survey.

Teachers’ comments about the Moderation Day discussed the importance of connecting with other teachers, the usefulness of professional learning that focused explicitly on real and relevant student work samples and examples of their own teacher practice, and the utility of focused time to work together with other teachers on issues related to assessment and pedagogy. The teachers’ comments were generally very positive and called for similar moderation processes in the future.

In relation to principles of effective moderation, the Year 1 Checkpoints Assessments moderation processes provided those engaged with the Year 1 Checkpoints the opportunity to:

- Discuss their judgments made about student achievement of particular year 1 Literacy and Numeracy Indicators – supporting consistency and reliability;
- Collaborate through a process of sharing work samples and evidence of assessment practices in a way that promotes the development of shared understandings and language – supporting collaborative and shared planning;
- Be presented with a clear and simple approach to making consistent judgments about student achievement. This was supported by the
materials provided to facilitators and the link to real work samples as the key focus of the moderation day – supporting the development of clear and simple standards and expectations; and

- Focus on a broad range of student achievement. This needs to be developed in a way that moves beyond deficit explanations of student performance, particularly as it relates to cohorts of students – supporting high equity approaches.

It is clear that future development of the Year 1 Checkpoints should include a moderation process. A consideration for those involved in the future development of these materials and processes is that in the 2010 Supported Trial, the Moderation Day served as a professional development opportunity that helped to focus the trial more generally. In future iterations of the implementation of the Year 1 Checkpoints if the focus shifts to account for relevant systems’ recording and reporting requirements then the opportunities afforded by the Moderation Day in 2010 will need to be provided in other forums.
6. Issues for Applications of Early Years Assessment within a Balanced System-Wide Approach to Assessment

Introduction
This report supports the development of teacher capacity in assessment. It supports the achievement of a system-based approach to assessment that is culturally fair, rigorous, and that offers a variety of opportunities for students to be assessed, for different purposes, and through different means. With the key purposes of assessment in schools being to improve teaching and learning, and to provide meaningful data and reports on student achievement to students, teachers, schools, parents and community as well as systems, assessment should be linked to initiatives in curriculum and pedagogy. Assessment should be educationally useful, and as such Queensland’s early years approach to assessment in literacy and numeracy must continue to have a focus on informing future teaching and learning decisions for all students.

Taken alone, current approaches to national testing are problematic in that they can only ever be a point-in-time snapshot of students’ performance on tests, limited in their validity and reliability by the very nature of their national or standardised status. The fact that we are currently operating in a context where the value of national testing as accountability is unquestioned, means that a balanced assessment program that surrounds the national testing initiatives is more important than ever. Part of this balanced approach must be an early years literacy and numeracy assessment tool that is rigorously valid and reliable, and culture fair and inclusive.

The purpose of this evaluation was to gather information about the appropriateness, effectiveness and utility of the June, October and February (trialed with prep teachers and students in November) Year 1 Checkpoints Assessments, and to observe and gather data at moderation sessions after the October Checkpoint Assessment.

Our approach to the task detailed in this report has been simple:

- We reviewed the literature field on early years assessment, and moderation processes. This framed our examination of the assessment tools and more broadly the current context of early years assessment within Queensland;
- We reviewed approaches taken to early years assessment in literacy and numeracy by other Australian systems, and revisited the past critiques of the Year 2 Diagnostic Net;
• We surveyed teachers and school-based administrators who were participating in the Supported Trial to gain information about their opinions of the Year 1 Checkpoints as assessment tools;

• We interviewed selected teachers to gather information on how the Year 1 Checkpoints were used and issues related to this in schools. To this interview data we added the range of qualitative comments taken from the surveys;

• We attended Supported Trial events and activities and collected observations of activities, and teachers’ oral and written comments; and

• We reviewed the materials, accessing expert opinion\(^6\) to support this review in order to comment on their purposes and fit to the task.

In this final section, we offer a series of findings and recommendations broadly contextualised in the context of the early years of schooling in Queensland.

**A Systems-Based Approach to Early Years Assessment**

The Year 1 Checkpoints, their purposes, effectiveness and quality need to be considered within the broader context of assessment in the first years of school. Any decisions made about the use of the Year 1 Checkpoints will have implications for other elements of this system-based approach. Consequently as decisions are made about the Year 1 Checkpoints and the purposes to which they will be deployed, consideration of the assessment landscape in the early years of schooling is paramount.

The Year 1 Checkpoints enter an early years landscape where there is mixed uptake of system initiatives and variable assessment capacity. Several systems initiatives over the past decade have focused on supporting and building teacher capacity in assessment in the later years of primary and middle schooling, and the senior schooling moderated school-based assessment processes have been recognised as supporting teachers in developing assessment capacity over many years. In the early years of schooling the Year 2 Diagnostic Net has remained the centre of systems-based assessment initiatives in the early years of school. It is now sixteen years since the net processes were first implemented and during that time there have been few changes to the materials and processes of assessment.

---

\(^6\) This review was supported by the expert comments of Karen Dooley (QUT) and Beryl Exley (QUT). Their input to the analysis of the materials presented in this report is acknowledged here.
The early years of schooling already have several components of assessment that are in place. It is important to ensure that decisions about the inclusions of a new tool be made in light of these established measures.

**Positioning Prep and the ELR**

The fact that the Prep year is not a mandatory school year has implications for the involvement of this year in a systems-based suit of assessment. The Early Learning Record requirements have supported the collection of ongoing, purposeful work samples and the development of assessment techniques that are linked to teaching and learning in the Prep year. As Prep teachers gear up to be involved in the Australian Curriculum implementation, the first time they have been expected to participate in a P-10 curriculum in Queensland, finding ways to link the assessment work of Prep teachers and students to the curriculum while keeping the principals of effective early years assessment and pedagogy to the forefront will become more important. The stakes for Prep are being raised. The placement of a February Checkpoint in the beginning of Year 1 is with good reason and we are not recommending a shift in timing. However we do recommend that the Early Learning Record and the Prep Indicators be considered in future decisions about the Year 1 Checkpoints and their use within the full suite of early years assessment in a Queensland system-based approach.

**The Year 2 Diagnostic Net**

Calls for reform of the Year 2 Diagnostic Net have been issued across a considerable time period and these have been detailed earlier in this report. However there have been significant positive effects of the Net for teachers and students in the early years of schooling also. It has provided a focus on and for assessment of students’ achievement in literacy and numeracy early in students’ schooling careers. Prior to its inception there was no such instrument and indeed no such focus in the early years of schooling in Queensland. Beyond providing opportunities to monitor student progress and procedures to validate teachers’ judgments the Net has also:

- Acted as a mechanism for professional development of teachers in issues related to assessment, literacy and numeracy, and learning;
- Provided teachers across the first three years of schooling with a common language to talk about students, learning, and achievement;
- Encouraged the sharing of ideas and expertise across schools and clusters by supporting Key Teachers and a moderation process for teachers involved in the validation processes of the Net;
- Provided a framework for matching teaching to assessment, at least for selected students;
- Provided the platform for additional funding to be provided to schools to support targeted students learning needs early in their school lives; and
Provided a path for expert literacy and numeracy teachers (Key Teachers) to take up extra responsibility without necessarily leaving work in classrooms.

However there remain critical issues with the Year 2 Diagnostic Net that should not be ignored in a renewal or review process of early years' assessment in literacy and numeracy. These concerns have been detailed in at least three reports, and have been summarized in this report also. The purpose of this report has not been to evaluate the Year 2 Diagnostic Net, nor to propose an alternative to the Net, and so this has not been part of the work completed. Importantly for the work that follows the delivery of this report, as decisions are made about the deployment of the Year 1 Checkpoints, the critiques offered over time and context about the Net and its processes should be considered.

National Assessment Program – Literacy and Numeracy (NAPLAN)

Year 3 is the first juncture where students' achievement is assessed using mandated common tests of literacy and numeracy. Since the introduction of the National Assessment Program – Literacy and Numeracy in 2008, students in Years 3, 5, 7 & 9 have been tested in May of each year. Data is reported to systems, schools, parents/caregivers (about the performance of individual students) and the public more generally (about the performance of schools and systems). The program replaced Queensland’s State-wide Tests in Aspects of Literacy and Numeracy, used prior to NAPLAN’s introduction to collect information about students’ achievement in literacy and numeracy in Years 3, 5, and 7. The implications of such a testing program, for individuals, schools and systems have been detailed extensively elsewhere and will not be reiterated here. However the inclusion of standardised tests in the early years of school raises the importance of including these tests in a balanced system-based approach to early years’ assessment in literacy and numeracy. This is particularly the case as these tests become more high-stakes for individuals, schools and communities.

The Year 1 Checkpoints Assessments in Literacy and Numeracy have been designed as a way to promote effective assessment of literacy and numeracy in Year 1. Their deployment must take account of the landscape that they enter. They will not stand alone as an initiative, instead becoming part of the initiatives – at a system, school and class level – that Year 1 teachers negotiate as they assess, plan and teach with their students.

The Year 1 Checkpoints aim to focus teachers’ attention on assessing students’ achievement in order to inform future teaching and learning. They focus on teacher judgment of students’ achievement of targeted Year 1 Indicators through the collection, analysis and reporting of evidence produced as part of everyday teaching and learning activities. The components of the Year 1 Checkpoints Assessments which are being suggested as effective assessment practices for Year 1 include:
The **February Checkpoints Assessments** are a list of assessment opportunities rather than prescribed assessment activities. As such within these instruments there is some flexibility for teachers to cater to the needs of diverse student populations. There is a good range of assessment types and methods in the suggestions. The February Checkpoints focus on what students can do rather than what they cannot do, so a deeper level of diagnostic analysis of individual students’ needs relies upon the professionalism and capacity of the teacher. The suggested contexts for assessment are clearly linked to the framework being utilised with targeted indicators clearly stated. However this is the only structure provided for understandings of the models of literacy and numeracy being deployed.

The absence of any implementation advice, strategies, links to teaching strategies or support to teachers in using the evidence collected to inform teaching means that the diagnostic power of the February Checkpoints relies on pre-existing capacities of teachers in an absolute sense. The Year 1 Checkpoints Assessments aim to support and develop teacher capacity in assessment in the early years of schooling, and develop effective practices of assessment for learning – that is for using assessment to inform future teaching and learning – but this would not necessarily be achieved through the February Checkpoints without the provision of relevant support materials and focus.

**The June Checkpoints** involve three literacy and four numeracy Assessments covering aspects of reading and viewing, writing and designing in literacy and number, measurement, space and algebra in numeracy. The focus of these assessments is variable, as is the clarity with which they achieve a clear link between the assessment item and what is being assessed.

**The October** are written as assessment units and as such fit well within common early years pedagogy and assessment practice. Ensuring the approaches taken within each unit are based in sound literacy and numeracy understandings – and justifying the choices made to teachers is important. Issues related to providing fair and accurate information about all students could be addressed in part by the provision of other alternative assessments.

**The support materials** provided were well received by teachers and supported the materials effectively. There is a need for further front loading materials to address some of the issues with clarity of links and purpose, validity, and issues related to differentiation of assessment to be addressed.

**The processes of moderation** provided teachers with a forum for professional development and confirmation of teacher judgments.
The principles and focus of the Year 1 Checkpoints Assessments is educationally useful. Issues related to time and workload require consideration in any further development of the materials. There has been much feed forward provided in this report in order to support the future development of the Year 1 Checkpoints. If this feed forward information is taken into account in the continued development of the Year 1 Checkpoints then the approach taken by the Year 1 Checkpoints Assessments is one that could build teacher assessment capacity, inform quality teaching and learning for all Year 1 students, and support the improvement of outcomes of school literacy and numeracy.

As such the Year 1 Checkpoints could be a useful tool for use by Year 1 teachers to assess, analyse student achievement, and plan effective future teaching and learning for year 1 students. With issues related to differentiation of assessment and culture fair assessment addressed, the Year 1 Checkpoints have the potential to provide a focus on quality assessment, teaching and learning for all students. Their links to a framework of end of year expectations provide a focus on teaching and learning, rather than normative descriptions of individual capacity and development. They could help to frame an approach to quality assessment practice in year 1 and be positioned within a systems-based approach to quality assessment for a variety of purposes across P-3.

If deployed to a task beyond these limits, there are implications of this institutionalisation process that must be considered in that event.
8. Conclusion

Our discussions with teachers and school-based staff working in the early years has uncovered a map of variable engagement with assessment in the early years, and specifically Year 1. Many teachers and schools had well developed, and efficient systems of assessment that worked to link quality assessment with curriculum and pedagogy in the interests of all their students. However this was not always the case.

There was evidence that the Year 1 Checkpoints had focused teacher’s attention on assessment and how it related to quality teaching and learning. There are instances in the data that point clearly to the fact that the focus on making judgments about the achievement of all students had focused literacy and numeracy teaching in educationally useful ways.

However there is also evidence that teachers struggled under time constraints and the workload attached to assessment practices within some of the tasks. The review of the materials raised issues related to the Year 1 Checkpoints and the Year 1 Indicators and the clarity with which they expressed their foundations in current understandings of literacy and numeracy, purposes, intents, and approaches to dealing with differentiation of assessment in the interests of fair and equitable assessment for all students.

With these issues in mind and with continued development, the Year 1 Checkpoints have the potential to focus assessment in Year 1 and to support the link between assessment and future planning, teaching and learning. As such they have the potential to support and develop teacher assessment capacity, and take a position within a systems-based approach to quality early years assessment.

Any deployment beyond this purpose would have subsequent consequences for the instruments, their validity, reliability and the educational usefulness to which they can claim.

Findings

F1. That all Australian State/Territory systems are currently engaging their early years’ teachers and students in some form of assessment in literacy and numeracy. These tools vary in content, emphasis, theoretical foundations and quality. The expectations of teachers and their reporting mandate also vary. The foundation of Qld’s current approach is the Year 2 Diagnostic Net, which has been in service for sixteen years, despite numerous calls for its renewal or replacement.

F2. That the P-3 Indicators state expected learning for each year level and thus differ from taxonomies of literacy and numeracy development such as the Year 2 Diagnostic net continua. This difference is related to framing much more than it is to content.
F 3. That there is evidence that many teachers had not used the Indicators as a framework prior to being involved in the Supported Trial. Some teachers had difficulty understanding the format, content, purposes and intent of the Indicators.

F4. That there is evidence that the Year 1 Checkpoints were generally well received by many teachers and school-based staff as items of assessment for the early years. This is not to say that teachers have not requested and demanded many modifications. These modifications were based on: the call for correction of errors; difficulties with understanding the purposes of individual assessment requirements; difficulties with the format or structure of specific assessments (particularly issues with a bias toward one-to-one individual assessment); issues related to making amendments to the assessments for specific students or cohorts of students; issues related to the quantity of assessment items and issues of time and workload; and issues of timing, time and workload more generally.

F5. That the additional support materials were generally well received by teachers. That interactive or digital versions of some of these support materials – particularly the DAAR – would support teachers in their use of these materials.

F6. That the expertise and professionalism of the QSA team supported teachers and administrators in their assessment work in schools. That there needs to be a variety of approaches to training and advice giving so as to accommodate teachers in a variety of contexts and the local constraints of these contexts (such as connectivity issues), and a variety of levels of teacher understanding, experience and IT capabilities.

F7. That, contingent on the suggested changes from chapter 4 of this report and the achievement of F8 and F9 below, the Year 1 Checkpoint Assessments generally follow the principles of quality early years assessment for learning. As such the materials could be used to support year 1 assessment and subsequent planning and teaching. They have not been Evaluated in any other capacity.

F8. Issues related to cultural fairness of the Year 1 Checkpoints have not been adequately addressed to date.

F9. Issues related to the capacity of the Year 1 Checkpoints to deal productively with diversity of experience and specific needs of students, and issues related to building teacher capacity in differentiation of assessment items to ensure the processes of assessment are based in equity and justice have not been adequately addressed to date. Teachers reported making amendments to the assessment processes to accommodate the needs of a variety of students, but there was no advice provided about such differentiation.
F10. Early years teachers in Queensland have focused on validating children within the Year 2 Diagnostic Net processes. Despite whether this was the original intent of the Year 2 Diagnostic Net materials, there is much evidence to support that this is the case. Linking the Net validation tasks with identification of individual students for funding purposes has ensured that the assessment items be understood in this way. The Year 1 Checkpoint materials and moderation processes have the potential to refocus teachers toward using checkpoints to validate teacher judgments rather than validate children. This has the potential to focus assessment on future teaching and learning.

F11. That the social moderation process implemented as part of the Supported Trial supported consistent teacher judgments of students’ work in Year 1. As a professional development opportunity teachers rated the moderation process as useful. Moderation has the potential to support a focus on validating teacher judgments.

F12. That the Year 1 Checkpoint Assessments currently provide a framework and a suite of assessment instruments for use in monitoring students’ achievement in literacy and numeracy in Year 1. There are issues that need to be addressed with individual Assessments and with the package as a whole (see chapter 4). The Year 1 Checkpoints can currently be used at three points across the year to monitor students’ achievement toward end of year indicators in literacy and numeracy so as to inform future teaching and learning. They do not currently feature expectations, standards, and resourcing decisions that would allow them to be deployed beyond this purpose. Their link to an end of year expectation has implications for their deployment as a diagnostic instrument to identify students for intervention and/or funding.

Recommendations

R1. That Queensland systems work toward a consistent, quality and rigorous system-based approach to early years assessment and diagnosis of literacy and numeracy learning needs across P-3 as a matter of urgency. In the context of education reform in Australia currently, it is important that Queensland have a current and well-resourced, high quality, high equity approach to assessment in literacy and numeracy in the early years of schooling. These decisions must take account of the measures already in place and their utility, as well as teacher workload, and the consequent delivery standards required to support a cohesive approach across P-3.

R2. That the QSA and/or EQ establish a working group to oversee the progress of a system-based approach to early years’ assessment in literacy and numeracy. The task of this working group would be to design and develop a coherent approach to early years’ assessment for
Queensland schools. The group would require technical expertise in measurement, evaluation and assessment, along with expertise in current approaches to literacy and numeracy, Indigenous education, and early childhood education. Their purpose would be to achieve a state-wide approach to early years’ assessment that is coherent, culturally fair, rigorous, valid and technically advanced, as well as realistic in terms of teacher workload and the delivery standards required for implementation.

R3. That any decisions about the uses and purposes of the Year 1 Checkpoints be made in light of the findings of past reviews of diagnostic assessment in Queensland. That the calls for culture fairness, validity, reliability and equity be considered as paramount as assessment instruments and frameworks are put to work as part of a system-based approach to assessment of literacy and numeracy in the early years.

R4. That a review of the content of the Literacy and Numeracy Indicators as they relate to current understandings of literacy and numeracy, and current and future curriculum expectations of literacy and numeracy across the curriculum be conducted.

R5. That appropriate use of language to describe the Indicators as expectations of learning and not a map of student development be used. That relevant teachers and school administrators receive advice and training which includes opportunities to develop an understanding of the underlying theoretical foundations of the P-3(9) Indicators as a frame for assessment purposes.

R6. That if the Year 1 Checkpoints are to be used in Queensland schools, that relevant teachers and school administrators receive advice and training on the Year 1 Checkpoints, the principles of literacy and numeracy teaching underlying the materials and process, the principles of assessment for learning, achieving consistent teacher judgments, the use of assessment for future planning and teaching, and processes of moderation for validating teacher judgments.

R7. That if the Year 1 Checkpoints are to be used in Queensland schools, that the introductory statement of the Year 1 Checkpoints be extended to address not just principles of assessment, but also the theoretical foundations of literacy and numeracy being deployed in the Year 1 Checkpoints, and advice on how the Indicators and Checkpoints are linked.

R8. That if the Year 1 Checkpoints are to be used in Queensland schools, that the advice given in Chapter 4 of this report be used to guide modifications. And that the QSA consider ways to simplify the assessment items as they are modified and alternatives are designed. This does not suggest a lowering of standards, but rather a
consideration of issues of classroom implementation and organisation, teacher and student assessment workload and time, content and construct validity, issues of language and content, and the dynamics of individual (one-to-one) and group assessment in the early years of schooling.

R9. That if the Year 1 Checkpoints are to be used in Queensland schools, that the QSA and EQ develop and trial alternative instruments as they modify and further develop the suite of materials so that teachers and schools eventually be provided with a bank of assessments from which they are able to select instruments that work within their site-specific contexts.

R10. That this development process (R9) include the development and trial of culturally relevant materials for Indigenous students across a variety of contexts. That Indigenous expert educators, with a broad understanding of the particularities of location evident for some students in Queensland schools and technical assessment and literacy and numeracy expertise drive this design and development process.

R11. That teachers in the early years receive advice and training in approaches to assessment that are culturally fair and able to deal with diversity and specific student learning needs. That this advice and training also focus on how best to record, monitor, analyse and report the differences that students bring to assessment tasks in educationally useful ways.

R12. That a process of social moderation be implemented as part of any implementation of the Year 1 Checkpoints.

R13. That issues related to delivery standards, such as required funding for the assessment process, be considered as part of any implementation plans of the Year 1 Checkpoints.

R14. That positioning of the Year 1 Checkpoints in relation to the P-3 years of schooling be considered as part of any implementation plans of the Year 1 Checkpoints.
References


Murphy, J. (no date). Teacher judgment in the application of standards in high stakes assessment: How moderation is interactionally accomplished. Queensland Studies Authority.


**State and Territory Websites**

Australian Capital Territory

New South Wales

Northern Territory
[www.det.nt.gov.au](http://www.det.nt.gov.au)

South Australia

Tasmania
[www.education.tas.gov.au](http://www.education.tas.gov.au)

Victoria

Western Australia
[www.det.wa.edu.au](http://www.det.wa.edu.au)