

Executive Summary

The National Accelerated Literacy Program (NALP) in the Northern Territory (NT) is an ambitious attempt to implement a method of literacy teaching, known as Accelerated Literacy (AL), in 100 primary and secondary schools in the NT. NALP attempts to address the literacy outcomes of the NT's Aboriginal population and is the largest project of its kind in the NT.

This evaluation of NALP in the NT was carried out by the School for Social and Policy Research, Charles Darwin University (CDU). With some variation in timing of data collection for specific elements of the investigation, it covers the period from the program's inception in 2004 to early 2008. The report has two volumes. Findings, discussion and recommendations are presented in Volume 1, while methodology and supplementary material are documented in Volume 2.

The evaluation examines two areas of critical concern:

1. How was the program's implementation in schools regarded by participants and did it lead to the kinds of teaching required by the Accelerated Literacy method?
2. What are the student learning outcomes of the program as measured by the available assessment tools?

To answer these two questions, a mixed-methods approach was adopted. The views of teachers, coordinators and others on the implementation of AL were systematically canvassed by means of focus group interviews and a survey of all teachers and coordinators. A valid and reliable method for observing teaching practices in classrooms was developed to assess the degree to which the pedagogy had been effectively transferred to practice. Both survey questionnaire and observational protocol were designed after an extensive review of literature, examination of videos of AL practice and a series of formal consultations with AL developers, expert AL coordinators and managers.

To answer the second question, data were provided on a range of variables relating to student assessment outcomes and characteristics of schools, teachers, students and locations and subjected to statistical analysis to ascertain rates of accelerated progress and to identify determinants of learning outcomes of students. The development of the methods, reliability testing and training for both components of the project is fully described in Volume 2 of this report.

Data-gathering, scope and limitations

The three major elements of this evaluation include a survey of practitioners supplemented by focus group interviews; an observational study of teaching practices in a sample of participating schools, and a multilevel statistical analysis of system-wide student assessment outcomes. The timeframes for data-gathering for these three elements of the evaluation are:

Survey (online and hard-copy) and interviews: January - April, 2008

Observational study (including pilots): April 2007 – April 2008

Multilevel statistical analysis dataset: January 2005 – December 2007.

In considering the outcomes reported here, the reader should bear in mind the scope of NALP in the NT and the fact that many, if not the majority of schools participating in this evaluation have only been recruited to the program at various times within the last four years, with PD workshops, materials and resources developed, written and disseminated as they became available to coordinators already operating in schools. This evaluation therefore in effect describes the results of ongoing capacity-building within the schools and systems working to implement Accelerated Literacy to the beginning of 2008, rather than the outcomes of an already established system.

Implementation Outcomes: Professional Development of AL Staff

Survey responses were received from 259 teachers, ALSBCs and principals. Focus group interviews were conducted with 211 teachers, school-based coordinators and members of the DET AL team.

Levels of professional development and support received by AL teachers and ALSBCs were established from responses to the survey. PD support provided within NALP (by both CDU and DET personnel) occurs both at formal PD workshops (referred to as PDs or PD workshops); and in-school PD support provided by the DET AL coordinators (according to the school's phase of implementation). In this report, the term *in-school professional development support* or *in-school PD support* will refer to the latter kind of school-based support provided by visiting DET AL coordinators to ALSBCs and AL teachers. As reported by practitioners in response to the survey, *total hours of PD support* received refers to both PD workshops and in-school PD support. Support is also provided to teachers by ALSBCs.

There were significant differences in attendance of the PD workshops, with a decline in participation from PD 1 (attended by 82% of practitioners) to PD 2 (61%), PD3 (43%) and a fourth Assessment Workshop attended by only 13% of teachers and by 16% of practitioners overall. Overall, attendance of ALSBCs at PD workshops was higher than others, for example 35% attending the Assessment Workshop. The mean numbers of workshops attended were 1.97 for teachers, 2.5 for ALSBCs and 1.52 for principals.

Survey data were analysed according to the remoteness of school location. Practitioners in Alice Springs and Katherine had significantly longer experience teaching AL, and had been employed at their current school for significantly longer than teachers elsewhere. Overall, close to 50% of all survey respondents had one year or less of AL teaching experience and 33% at the time of survey were in their first term at their present school. However, in very remote schools teacher tenure was significantly shorter, on average less than 8 school terms compared to over 14 terms in Alice Springs and Katherine. For PD1, attendance of Darwin practitioners was somewhat lower than the average at 72%, with 89% of practitioners in remote schools (Katherine and Alice Springs) and 85% in very remote locations (Tennant Creek and all other remote community Community Education Centres [CECs]) attending PD1. Alice Springs and Katherine had the best attendance at PD workshops overall, with a mean of 2.36 attended; Darwin the lowest at 1.93 attended; while practitioners in very remote schools attended 2.03 workshops.

Schools in very remote locations had reportedly received significantly less total support, including in-school PD support (provided by the DET AL team) than teachers in the other locations, with a mean of 21.91 hours, compared with means of above 30 hours for the other schools.

Analysis of both focus group and survey findings sought to identify the views of practitioners about the effectiveness of the implementation program and their relationship, if any, to the levels of PD and support received.

Teachers were significantly less positive in their rating of AL and almost all facets of the implementation than either ALSBCs or principals. In general terms, teachers gave a neutral rating of the degree to which training and support left them confident to teach AL and all elements of the AL teaching sequence. The highest approval was reserved for the effectiveness of the DET AL team, and the helpfulness and availability of program materials and resources. The lowest rating was given to assessment of students and the usefulness of assessment information.

Teachers gave a neutral rating to the performance of ALSBCs and in focus group interviews gave a mix of opinions of their effectiveness ranging from highly critical, with very low evaluations of their performance, to a few positive examples of effective ALSBC support. A number of ALSBCs complained of a lack of training and confidence in their role. Principals, by contrast, appeared to significantly overestimate the effectiveness of the ALSBCs by comparison with teachers.

In terms of the relationship between levels of training and support and practitioners' opinions of AL and its implementation, analysis found no relationship to the number of formal workshops attended. However, there was a significant association with the amount of total PD support received including workshops and in-school support. This suggests that in-school PD support at the current stage of implementation is an important factor in securing practitioner commitment and motivation. This was generally confirmed by focus group interviews, in which many teachers were critical of PD workshops, and were either more positive about the visiting support provided or were of the view that there needed to be effective follow-up and on-site practical demonstration of PD workshop content to be able to implement AL with confidence.

AL in the Classroom: Observed Teaching Practice

For the observational study, 68 teachers teaching 183 AL lessons in 36 schools were observed. The sample was largely determined by the number of schools available to participate in the study at the time. Teachers' and ALSBCs' survey responses were analysed to identify characteristics of the sample in relation to observed AL teaching performance. Characteristics of teachers observed compared with the full survey sample indicated that observed teachers were less experienced, but had attended significantly more PD workshops than the full survey sample. They were not different in terms of total levels of PD support received, time at present school and remoteness (teachers in very remote schools were two-thirds of the sample).

Teachers' classroom teaching practices were assessed using a purpose-designed observational protocol administered by a trained team of field researchers. A high degree of inter-rater reliability was established after thorough training and testing at pilot. This protocol rates compliance with the AL method on nine dimensions of the AL teaching sequence. Practices were rated on a five-point scale. The middle score of 2 was required to meet the criterion "effective implementation" of AL practices. Teachers were also assessed on the frequency of literacy teaching practices on six additional scales derived from recent research on best practice in literacy teaching (Louden et al., 2005).

Of the nine dimensions of AL practice, only one element, *whole lesson organisation*, was rated by observers at above the criterion level of 2.0, "effectively implemented". All others were clearly below criterion level. The dimension, *whole lesson interaction*, was rated at approximately 1.7 (with the upper margin for error just under 1.9). The remaining seven dimensions (including items of writing, spelling, high and low order orientation, etc.) were rated much lower at 1.5 or lower, that is, hovered at or just above the rating "attempted but not effectively implemented". Writing-related elements of the sequence, although expected, were observed in very few cases at all.

Whole lesson organisation refers only to use of appropriate materials and physical organisation of the classroom. This shows that teachers had been able to select texts, deploy resources and organise the lesson around the key elements of the teaching sequence. However, the scores on other elements of the sequence show that they had not been able to effectively implement the key strategies of the AL teaching sequence. The frequency of "differentiation" achieved was particularly low. This suggested that teachers, at present levels of training and experience, were struggling to reconcile the whole-class orientation of AL with teaching strategies oriented to the needs of individual students and smaller groups and that in general, higher-level AL skills (more consistent with a high frequency of differentiation) were not being achieved or integrated in teaching practice. Focus group findings clearly confirmed this interpretation. At interview, teachers directly indicated that they were struggling to incorporate specific responses to areas of need and that they found the whole of class approach of AL difficult to reconcile with an orientation to individual student needs.

Observed Practice, Professional Development and Support

Analysis of the respective contributions of training and support received, school location, and other factors to observed teaching performance remains somewhat tentative because of the small sample size of the study (n = 68). However, of the factors examined, there was a significant association between the level of reported in-school PD support provided (mainly by the visiting DET implementation team coordinators) and the observation of AL practice in classrooms. Regression analysis yielded the following findings:

1. The number of PD workshops attended did not predict observed teaching performance.
2. There is a significant association between the total number of hours of PD support provided by DET coordinators and observed teaching performance, accounting for 21% of variance in practices observed in the classroom.
3. At 50 hours of total PD support, teaching practice approaches the criterion level of “effective implementation of AL practice”.

Teachers in very remote schools reported receiving significantly less total PD support than other schools, in part as a result of higher teacher (and principal) turnover and lower levels of AL experience than schools in the major centres. Although not demonstrated with this sample, the performance of teachers in very remote schools (with some exceptions) may have suffered as a result.

The mean level of 37 hours of total PD support received by ALSBCs was sufficient for them to positively rate the implementation and methods of AL, but, according to observations, not to teach AL to desired standards themselves. Moreover, qualitative findings suggest that even when ALSBCs have received the 50 hours of total PD support required for effective teaching of AL, they are not necessarily able to effectively support teachers to teach AL. This has significant implications for the ability to embed AL sustainably at school level. It suggests that higher order training should be provided to ALSBCs to ensure that they have advanced skills and are not simply ahead of teachers by the number of hours of standard training received.

Teachers reported that the PD workshop program lacked quality control mechanisms, and that the absence of implementation checks to determine levels of teacher competence following both workshop PD and in-school PD support may have contributed to their lack of confidence regarding their understanding of the AL methodology and the skills required to implement the AL teaching sequence. Teachers suggested that PD3: Writing Workshop needs to be reviewed. Many teachers responded positively to a more recently introduced intensive PD workshop and were clearly more positive about the role of in-school PD support provided by the DET AL team. In-school PD support provides opportunities for DET coordinators to deliver elements of the content of the PD workshops and to assist teachers and ALSBCs in the practical application of methods.

In summary, the findings of survey, observations and focus groups are consistent with

the view that by early 2008 there had been an undertreatment of the population of teachers participating in NALP in terms of access to and timing of both workshop PD and in-school PD support, and that the contribution of the then current workshop PD relative to other forms of support could not be reliably demonstrated. The intensive workshop introduced in 2008 has not been evaluated. The apparent relationship between teacher-reported PD support and implementation outcomes, including the adherence to AL methods in the classroom, suggests that to achieve and sustain desirable thresholds of practice in schools, an optimum combination of PD workshop and targeted in-school PD support for teaching practice needs to be identified.

It may be a mistake to simply draw the conclusion that increasing the quantum or “dose” of visiting PD support to participating schools should be the sole response to the findings of this report. Firstly the level of exposure to formal PD must be increased. Secondly, given the interdependence of the theoretically oriented workshops and practical in-school PD support, any review of practical support should be carried out in conjunction with a review and evaluation of formal PD workshops, to ensure that the latter incorporate updated content, achieve higher levels of teacher satisfaction and can be more effectively translated into practice, including those areas of practice shown to be least adequately implemented. The significance of visiting in-school PD support by the DET AL team undoubtedly also reflects the fact that self-sustaining capacity had not been achieved at schools. Many teachers were emphatic that proactive school leadership was the most important ingredient of a successfully run program of AL teaching. Improved performance of ALSBCs and AL teachers and principals as leaders could eventually reduce the dependence on visiting support.

The various elements of PD, support, school capacity, leadership and other factors are therefore likely to interact in complex ways. Simple expansion of the quantum of support applied to schools may be an inadequate response (and may strain the resources of the reportedly already overstretched centrally located team). It is likely that, under present arrangements, support has not been as well targeted as it might have been, to ensure that schools with least capacity have had most support; it is possible that the reverse has occurred in some cases, with higher capacity schools receiving more in-school PD support. In considering future directions for NALP, priority may need to be given to identification of pre-existing school capacity before entry to the program in order to offset the demand for external support and improve prospects of rapid success. This may entail a review of the scale of the program.

Specific Areas of Concern

The evaluation identified some clear gaps in teaching performance. The current use of assessments by teachers to guide practice is inconsistent at best and many of the assessment tools provided are simply not being used. This applies generally, in that teachers report minimal participation in the administration of IL, ToRCH and WL assessments, as well as in the specific field of early childhood. Teachers report that there is inadequate feedback about results from ALSBCs and DET AL coordinators and that assessment data and system-reported data are not useful. Inconsistent use of assessments and the non-use of some assessment tools indicate that there is a need to review the place of assessment in AL training and practice and to strengthen the

guidelines for use of assessments by teachers.

The use of appropriate assessments may be particularly important for early childhood teachers and teachers of older non-readers. Based both on teacher testimony and on observation, there was little or no use of assessment for beginning readers and there are clear deficiencies in time allocation to instruction in spelling and writing. If the relevant assessments are not being used, then time given to teaching phonological awareness, letter/sound correspondence, decoding, spelling and writing is also likely to be inconsistent, with the result that insufficient time is being spent for students to achieve automaticity in these competencies.

Observational findings confirm the failure to reach criterion levels for the dimensions of writing and spelling and that most teachers are not spending sufficient time in the writing stage of the AL teaching sequence to ensure that students acquire writing skills. These findings are also consistent with focus group reports: many early childhood teachers at interview were particularly uncertain about the AL strategies for early literacy skills acquisition and about how to incorporate their existing teaching skills and training around phonological awareness, letter/sound knowledge, word attack skills and spelling, into the AL approach.

Levels of PD participation were very low in these areas and there appears to be a need to improve levels of training in practice requirements, to clarify roles and the allocation of time to these tasks. However, it is not likely that the lack of access to workshop training and in-school PD support fully accounts for the gaps and uncertainties in these important areas of teaching practice. Drawing on all findings, the conclusion of the evaluators is that the AL method should in fact be more explicit in its approach to early literacy development. AL theory is not sufficiently prescriptive in early skills acquisition, with the result that teachers must find their own way to incorporate appropriate methods into the AL approach. This may not be readily achievable for many teachers.

Student Outcomes 2005 to 2007

A systematic investigation of outcomes in all participating schools was conducted. This enabled identification of levels and predictors of accelerated reading gains of students who participated in NALP in the NT during the 2005 to 2007 period. Accelerated reading gain was defined as a student's rate of reading progress – *an annualised score greater than one reading year over four school terms* – according to either of two assessment measures: the Individual Level test (IL, n = 3,166 students), a purpose-designed observational test of reading accuracy, and the Test of Reading Comprehension (ToRCH, n = 941), by ACER (2003). Each assessment is applied at specific minimum entry levels: Transition level for IL; Year 4 for ToRCH. The analytical strategy was based on bivariate, multivariate, data-mining and multilevel procedures for estimating the effects of blocks of explanatory factors grouped by sequences, students, schools and regions.

Bifurcation of Outcomes in Two Samples of Students

Using two major instruments with different guidelines for measuring reading gain by students participating in AL, has meant that there are in effect two distinct but overlapping samples of students, one for each assessment type. Indigenous students attending schools in very remote communities were 57.3% (n = 1,657) of IL-assessed students, compared with only 9% (n = 144) in the ToRCH-assessed sample.

There was wide variation in average rates of accelerated progress for these two samples. The ToRCH samples outperformed the IL samples by .45 reading years per year with an overall rate of progress of 1.4 reading years per year, with 47.8% of students showing accelerated progress (progress greater than 1 reading year per year) and 24% showing no progress in reading scores. For the whole IL sample, the average rate of improvement was .95 reading years per year. In the IL sample, 31.8% of all students showed accelerated gains. However, only 19% of very remote Indigenous students recorded accelerated gains, as against an average of almost 46% for the other groups in the IL sample. In the IL sample, 38.9% of students recorded no progress in reading scores and most of these scored zero on both assessments.

Table 1: Characteristics of samples for IL and ToRCH assessments

	IL	ToRCH
By Size		
Valid sequences	6219	1573
No of students	3167	951
No of schools	52	32
By Socio Demographics		
Indigenous	80%	44%
Very remote	58%	15.4%
NESB (home language)	69%	38%
High school	14%	52%
Average school year level	4.84	6.73

The diversity of student characteristics for each assessment type thus coincided with a major bifurcation in outcomes that was particularly marked for IL assessments. The major explanatory factors for IL students were language other than English spoken at home and remoteness of region, with some influence of attendance rates, with Indigenous status less significant; while for ToRCH assessments, Indigenous status and language other than English spoken at home were the only significant effects, with no significant variation across regions. School differences, when these other factors were held constant, were very small for the IL, and not statistically significant at all for ToRCH assessments.

The low rates of accelerated progress for Indigenous students from very remote community schools on the IL test were associated with a much lower initial reading score (a mean of 1.85 reading years for the full IL sample, below 1 year for Indigenous very remote students, and just over 3 reading years for the other subsamples).

There was a very high proportion (53%) of average rates of zero progress across sequences and a strong relationship between lower initial or pre-program scores and overall rates of progress.

Explanation of Outcomes

There were considerable differences in the segments of the study samples defined by rates of acceleration, according to regional remoteness, Indigenous status and linguistic background. For example, 37% of female students aged 11 through 14 years in very remote regions achieved accelerated progress. This was well above average for the full IL sample and points to the possibility that there may be different growth rates at key periods of student learning in some groups. Indigenous students in Alice Springs and Katherine have higher average rates of acceleration than their non-indigenous peers, a finding favourable for the hypothesis of a catch-up of Indigenous and non-indigenous populations in these centres. Non-indigenous students in very remote regions have the highest rates of all six groups grouped by Indigenous status and region, outperforming their non-indigenous peers in Darwin schools. This works against the idea of remoteness by itself explaining poor outcomes and points to the need for further analysis of causes of difference between subpopulations, in this case, the distinctive characteristics or experiences of non-indigenous learners in remote communities.

When other explanatory variables are controlled for, it appears that a student's attendance rate does not have a strong or significant effect on average levels of accelerated progress. While attendance is important for the full IL sample, when between-school differences are controlled for, it has only a very small influence on acceleration and has no effect at all for the IL sample (excluding those students with a zero IL score on both assessments) or the ToRCH sample. This is an unexpected finding, since rates of attendance are commonly seen as a necessary, if not sufficient, condition for Indigenous literacy progress. While attendance is justifiably pursued as a general policy priority, the priority for AL should be a differentiated approach: to schools in terms of teacher training and adequacy of support, and to students on the basis of their reading level at point of entry to the AL program.

It must be noted that all of the influences in combination explain only a fraction (up to 20%) of the total variation in individual student rates of reading progress. This shows that most of the variation in the distribution of assessed rates of reading progress appears to be found in differences between students, teachers and classrooms. For this reason, and notwithstanding the need to target efforts to improve teaching performance, it would be misleading and unfair to use the existing outcomes measures as an indicator of the "performance" of schools or to rank schools according to their outcomes using these measures. Even with adjustments of scores for explanatory factors such as region and Indigenous status, any kind of comparison based on measured outcomes is of doubtful value for evaluation, for policy and management, or for feedback to practitioners. Performance of schools should focus on measures of leadership, capacity, resource allocation, training, and the quality of AL teaching.

Indigenous Status and its Correlates

The findings suggest that Indigenous status as an explanatory factor needs to be put into a wider causal context. For both IL and ToRCH assessments, there is little support for the case that negative effects of Indigenous status on reading progress can be explained solely in terms of remoteness or regional disadvantage.

The effects on accelerated progress rates for Indigenous status and remoteness predictors are expressed in different causal modalities. This analysis shows that the most powerful effects determining poor rates of progress are a combination of “language other than English spoken at home” and “reading age at first assessment”, with both concentrated among the population of Indigenous students in very remote communities. In short, they suggest heterogeneity in Indigenous populations according to linguistic, cultural and other factors that are relevant to student learning. This analysis suggests that, for the IL sample, there are some small differences associated with remoteness of school, and the findings of other sections of this evaluation indicate that very remote schools experience lower levels of total PD support and higher rates of teacher turnover. However, these factors almost certainly do not account for the differences in outcomes between very remote Indigenous students and other groups.

Lower Initial Reading Score and Very Remote Students

The low rates of accelerated progress for the Indigenous students from very remote community schools on the IL test were associated with a low initial reading age score of below Year 1 reading year level, compared with just over Year 3 reading year level for the other subsamples. This group also has a very high proportion (53%) of average rates of zero progress across sequences and a strong relationship between lower initial or pre-program scores and overall rates of progress. The factor of reading age interacts with the other factors outlined above to produce a very powerful effect, with the lowest rates of progress among Indigenous students from very remote backgrounds who are younger and have a lower initial reading age at first assessment. The analysis of reading age, school year level and rates of accelerated progress shows that a gap of about three reading years is established between very remote Indigenous AL students and all other AL students by about the third school year and that this gap (both in reading years and in proportions showing accelerated learning) persists through middle school. Although there is evidence that the gap in proportions showing accelerated progress is less pronounced between this group of AL students and others by about Year 7, there is not sufficient improvement to close the gap between outcomes for very remote Indigenous students and the remainder of the student population.

Because of the lack of information provided by the IL assessment for the large number of non-readers, supplementary analysis of data from five AL schools (two provincial/urban, three very remote) was conducted using the LLANS assessment instrument (Australian Council of Education Research, 1999). This study indicated that just under 40% of students who appeared as “non-improvers” on the IL assessment (88% of whom were pre-scale on all assessments), showed at least some improvement when assessed by LLANS. However, the students re-classified as “improvers” under the LLANS were concentrated in the two urban schools compared

with much lower rates in the remote schools. The results thus strongly suggest that a switch to the use of a more sensitive evaluative tool such as LLANS for NALP-related progress would be likely to accentuate, rather than reduce, the literacy gap between urban and remote schools, as well as between very remote Indigenous and other students. It also confirms the significance of reading age at first assessment and the importance of the early years in establishing the gap between students. Nevertheless, for both evaluative and formative reasons, the analysis indicates that an additional or different mode of assessment for early readers and non-readers should be adopted by NALP.

Between Program Implementation and Learning Outcomes

Within the limits of the research design, it appears that NALP in the NT is delivering accelerated reading progress of greater than one reading year level per year according to benchmarked assessments for mainly urbanised students who have achieved at least NT Year 4 levels of reading performance. That is, there are positive signs that AL is able to contribute to closing the gap for Indigenous students in the major centres of Darwin, Alice Springs and Katherine. However, at the current levels of delivery of AL using current methods in NT schools, for most of those students who, because of cultural, linguistic and situational factors, have not yet established basic reading competence, the evidence suggests that Accelerated Literacy has yet to demonstrate the achievement of substantial accelerated gain.

These findings must be considered against the findings of direct classroom observation, surveys and interviews. They show that, notwithstanding the progress in implementation of the program in NT schools, the level and quality of implementation achieved by the end of 2007 had not been sufficient for by far the majority of ALSBCs and teachers to achieve standards of teaching consistent with the requirements of the AL method. This reflects the need to strengthen access to and participation in both workshop PD and in-school PD support with the aim of improving the local capacity of schools to sustain the AL method. The continuing centrality of the DET AL team to the program in its current state of development reflects the fact that the transfer of necessary capacity to schools is yet to be achieved.

The findings suggest that in the next stage of implementation, NALP should address the bifurcation in outcomes between urban and remote contexts. The outcomes suggest not only that the quantity of training and support may need to be improved, but that the AL method needs review and further development for important target groups if it is to be effective for all students in the NT population. The retention of NALP as a population-level program for Indigenous students in very remote schools would be justifiable on the basis that existing capacity to deliver high quality, standardised training and teaching support should be retained and used and the program itself strengthened and evaluated to demonstrate its effectiveness for these populations.

In our view, and as outlined further below, the further development of AL PD and support needs to include: revision of and up-scaling of writing emphases; tools and development of letter/sound knowledge, word attack skills and spelling. In the absence of additional resources, the scale of implementation may need to be reduced

to enable selection of schools with the capacity to commit to intensive implementation of a revised program.

Evaluation for Sustainability

The limitations of the research design for this evaluation – specifically the absence of benchmark data and the inability to establish control or comparison groups – were largely a product of the scale and scope of the implementation program itself. This meant that such elements of design had to give way to the more urgent priorities of system- and capacity-building. However, it is now possible to see that further research can support ongoing improvement of the effectiveness of NALP. Mechanisms for continuous evaluation and quality improvement can be built into the program.

The evaluation has shown that maintenance of “program fidelity” can not be taken for granted, but that it needs to be periodically verified using appropriately designed methods. These methods should be adapted to support ongoing improvement of practice and to ensure that any innovations in teaching methodology are actually transferred to practice in the classroom. The teaching environment in which NALP operates is dynamic and changing; it is continuously subjected to the impacts of teacher turnover and other influences, which affect school capacity in the classroom. For this reason alone, a system to monitor and feedback practice development outcomes to schools and managers is desirable.

This evaluation has identified the need for NALP to incorporate a targeted program to develop and evaluate teaching strategies for readers with low initial reading scores, who in the NT are concentrated in remote locations characterised by distinctive ecological, cultural and linguistic circumstances. Any such program of research would collaborate with the ongoing implementation program, and include review and redesign of PD support, materials and resources and other mechanisms for guiding teacher practice.

We propose a two-tiered approach to program evaluation. The first tier would involve a cyclical evaluation of implementation both at school level and system-wide. It would aim to monitor achievement of designated elements of school participation and capacity, observe AL teaching practice in classrooms, and provide feedback of evaluation outcomes to participating schools and teachers. This component would be tailored according to the scope of the continuing implementation of NALP, and specific mechanisms and capacity requirements for schools to participate in the program.

The second tier of evaluation would directly target the bifurcation of student outcomes and would take the form of a targeted evaluation of a revised AL methodology. A formal review and evaluation of PD workshops and in-school follow-up, accompanied by review and redevelopment of AL tools and methods for early childhood and beginning readers, would precede the evaluation of revised second stage AL program in trial schools with comparison groups. Research-specific funding would be sought to augment resources available to this phase of the program.

Major Recommendations

These recommendations are intended to address the vulnerabilities of NALP and to improve prospects for effectiveness and sustainability as identified in the major findings. Our assumption is that to achieve the recommended changes and improvements within existing resources, some reduction in the existing or planned scale of the program is likely to be necessary.

Implementation and Sustainability

1. The NALP implementation strategy needs to address the bifurcation of learning outcomes in the population of participating students by specifically targeting the needs of Indigenous and very remote students and students with low assessed reading age.
2. The following broad strategies to improve implementation outcomes are recommended:
 - a) determine appropriate scale of implementation
 - b) select participants for a higher degree of school capacity to deliver AL
 - c) respond to regional differences in capacity to participate with appropriate levels of support
 - d) focus on building school-level capacity and commitment
3. It is recommended that DET and partners ensure that all AL teachers and AL school-based coordinators access indicated levels of PD workshops and other PD support. It is recommended:
 - a) that PD be accompanied by quality controls and follow-up by DET AL coordinators and PD developers
 - b) that participation in PD is linked with a cycle of evaluation of classroom teaching to support translation of PD into practice
 - c) that a regular cycle of refresher PD be designed and implemented
 - d) that ALSBCs receive advanced training in effective support for teaching
4. The system of in-school PD support should involve performance assessment and constructive feedback for teachers and coordinators and be implemented as an ongoing program to continuously develop professional standards.
5. The school phase model for implementation should be reviewed and linked to real and verifiable performance expectations and commitments, and be subject to audit in conjunction with an AL quality improvement program.

6. NT DET should review incentives and controls to encourage schools to meet program commitments; to maintain training levels, supports, processes and time-allocation; and to collaborate with assessment and evaluation processes.
7. A revised implementation strategy should target initial levels of reading competence and their social, linguistic and situational correlates. Key elements to be developed in conjunction with the evaluation program are:
 - a) that strategies for embedding assessment in teaching practice based on departmental requirements for compliance by teachers are implemented
 - b) that there be review of training, policies and support on assessment to support teaching practices, including those surrounding phonological awareness, letter/sound correspondence, decoding, spelling and writing
 - c) that the use of IL and ToRCH in conjunction with other assessments for beginning readers be reviewed for targeted subpopulations

Research and Evaluation

8. Continuing evaluation of classroom practices is recommended to ensure that required standards of teaching are achieved. For this purpose a model of evaluation to support continuous quality improvement in schools and clusters could be developed.
9. An evaluation program should be specifically designed and resourced to investigate the effectiveness of a revised AL teaching method for early childhood and beginning readers in selected very remote community schools
10. A population focus should be adopted for measurement of student outcomes across classrooms, teachers, schools, their types, sectors and locations, with appropriate analytical models developed for reporting outcomes, implementation measures and quality improvement goals.