Further Options for an Outcomes Framework for Reporting Purposes for the National Asian Languages and Studies in Australian Schools (NALSAS) Strategy

A paper prepared for the NALSAS Taskforce

Erebus Consulting Partners

February, 2001
Further Options for an Outcomes Framework for Reporting Purposes for the National Asian Languages and Studies in Australian Schools (NALSAS) Strategy

Table of Contents

1. INTRODUCTION ............................................................................................................................... 3
2. THE CONTEXT FOR OUTCOMES MEASUREMENT AND NATIONAL REPORTING IN AUSTRALIAN EDUCATION ........................................................................................................................................ 5
3. THE NALSAS STRATEGIC PLAN AND REPORTING AND ACCOUNTABILITY FRAMEWORK .......... 7
4. THE PLACE OF ASIAN STUDIES AND LANGUAGES IN STATE/TERRITORY CURRICULUM FRAMEWORKS ....................................................................................................................................... 9
5. CURRENT STATUS OF OUTCOMES MEASUREMENT OF ASIAN LANGUAGES IN STATES AND TERRITORIES ........................................................................................................................................ 16
6. POSSIBLE MEANS FOR ASSESSING AND REPORTING STUDENT OUTCOMES ........................................ 22
   Option 3 ........................................................................................................................................ 25
7. ISSUES IN DEVELOPING ASIAN LANGUAGES SAMPLE STUDY INSTRUMENTS ................................. 27
   7.1. Costs of a sample study ........................................................................................................ 32
   7.2. Who should be tested? ......................................................................................................... 27
   7.3. What should be tested? ....................................................................................................... 28
   7.4. When should students be tested? ........................................................................................ 31
Further Options for an Outcomes Framework for Reporting Purposes for the National Asian Languages and Studies in Australian Schools (NALSAS) Strategy

1. Introduction

This paper responds to a request from the Ministerial Council on Education, Employment and Youth Affairs (MCEETYA) NALSAS Taskforce, through the Commonwealth Department of Education, Training and Youth Affairs (DETYA) for the development of a position paper on further options for an outcomes framework for reporting purposes for the National Asian Languages and Studies in Australian Schools (NALSAS) Strategy. This strategy is a cooperative initiative of the Commonwealth and State and Territory governments, aimed at assisting schools to improve the participation and proficiency of student language learning in four priority Asian languages – Japanese, Chinese (Mandarin), Indonesian and Korean. The Strategy also supports the studies of Asia in the subjects studied by all students.

The NALSAS Strategy has been a significant initiative since its inception in 1994-1995. The resources committed to it by the Commonwealth, and by State and Territory governments have been substantial. The Commonwealth has provided over $144 million to support the study of Asian languages and studies of Asia since 1994-95, with funding in 2001 of $29 million. The goals for the original Strategy, as recommended in the 1994 Council of Australian Governments (COAG) report *Asian Languages and Australia’s Economic Future* were challenging:

“By the Year 2006:
- 60 per cent of students in Years 3 to 10 will be learning one of the four targeted languages;
- the remaining 40 per cent will be studying other languages
- 15 per cent of Year 12 students will be learning other languages; and
- all students in Years 3 to 10 will have Asia-content incorporated into the other subjects they are studying.”

An evaluation of languages teaching conducted in 1998 found that the indicators of success of the Strategy had been mixed. More importantly, the evaluation found that the quality of data available nationally about the status of language teaching and learning in schools was inadequate.

The NALSAS Strategic Plan for Phase 2 (1999-2002), developed by the Taskforce appointed by MCEETYA, proposed four focus areas for development:

---

Curriculum delivery
Teacher quality and supply
Strategic alliances
Outcomes and accountability.

This position paper addresses the last of these focus areas. The position paper is intended to be a further step towards establishing an appropriate outcomes reporting framework. The paper will also explore how outcomes reporting might best serve the information needs of the diverse stakeholders in the NALSAS strategy.

Some work has already been done in this area. NALSAS has funded a series of projects over the years that have attempted to identify “outcomes” of one kind or another. These projects have most often been concerned with the outcomes intended from language teaching and learning. The NALSAS Strategic Plan for Phase 2 sets out a framework for the collection of data about student participation in Asian language courses and other student learning outcomes in Studies of Asia. This will provide policy-makers with a much clearer understanding of the reach of Asian language programs nationally, particularly at the primary and junior secondary levels, where information is most lacking at present. What remains to be done is the development of an agreed framework for collecting data on student learning outcomes in Asian languages.

The term outcomes in this paper is used to mean the knowledge skills and understandings developed in students as a result of participation in language teaching and learning activities. Measurement of these outcomes necessarily involves capturing in some way the demonstrated proficiency of students in achieving the objectives of these teaching and learning activities. In assuming this definition, we are aware that there are other views about what outcomes are (this point is taken up again later in the discussion). This paper is not proposing to measure the outcomes of the NALSAS Strategic Plan per se, which legitimately has goals and objectives beyond the student domain. This paper concentrates most heavily on the Asian language learning aspects of the NALSAS Strategy, given that a separate baseline study of student knowledge and understanding of Asia has already been commissioned, which should provide appropriate data for this aspect of the Strategy.

In discussing the options for an outcomes reporting framework, this paper is primarily concerned with the mechanisms by which data about student learning outcomes can be collected and reported. The term “outcomes framework” is sometimes used by others to mean the intended or expected outcomes, or objectives for student learning that are contained in curriculum and syllabus document. There is an important distinction between this use of the term and how it is used in this paper. This paper does not attempt to define what the outcomes to be reported should be in any detail, (indeed this is a considerable piece of work in itself), but discusses some of the conceptual issues in doing so.

There are some significant issues to be resolved in undertaking such a task on a national basis. These include technical issues in measurement and calibration of language learning across different languages, practical issues of implementation of

---

2 Some examples here
data gathering and analysis including test development, test administration, and marking, and policy issues in reporting the outcomes of such testing. The paper will explore these, and other issues in more depth.

Resolution of these issues will require extensive negotiation between State/Territory and Commonwealth jurisdictions. These issues also touch on the work of other MCEETYA Taskforces, which will also need to be consulted before any measurement and reporting activities are implemented.

2. The context for outcomes measurement and national reporting in Australian education

Pressure for increased accountability has been felt in all areas of public endeavour for more than a decade in many countries around the world. The education sector in Australia has not been exempt from this trend. There are clear expectations from governments at all levels that recipients of public funds should be accountable for how those funds have been used, and more particularly, what has been achieved as a result of this expenditure. In relation to school education, the “results” that ultimately need to be considered are the knowledge, skills and values that students gain as a product of their schooling.

Changing attitudes towards accountability for outcomes by educational authorities can be seen in the increased willingness of systemic jurisdictions to publish data about student performance in ways which allow comparisons of various kinds, including comparisons over time, between geographic areas, between different demographic groups, and with standards or benchmarks. The level of detail now found in the Annual National Report on Schooling in Australia (ANR), or the Productivity Commission’s Report on Government Services would not have been acceptable just a few years ago.

However, while there may be general agreement in principle about the legitimacy and desirability of national reporting of student outcomes data, there is considerable disagreement about what should be measured, how it should be measured, and how it should be reported in any particular subject area. There are differences in opinion also about when particular assessments should take place, at what grade level(s), and under what circumstances.

The history of efforts to develop national benchmarks for Literacy is instructive in this regard. Debate about appropriate methodologies, standards and protocols has required several years of intensive work, and considerable resources to resolve. Even then, not all stakeholders are in complete agreement with the outcome, and there are those who suggest that the resources expended to produce nationally comparable data could be better spent on program delivery.

4 See for example, DETYA (1997). Mapping Literacy: Results of the 1996 National School English Literacy Survey Canberra: AGPS.
The States Grants (Primary and Secondary Education Assistance) Act 2000 provides a strong impetus for improving the status of student assessment and reporting processes, as a vehicle for understanding student outcomes in areas that have not previously been subjected to extensive public scrutiny, including Asian languages and studies of Asia. This Act requires that recipients of funds (such as NALSAS) account for expenditure of grants both financially and educationally. The Act states that educational accountability requirements for Commonwealth Schools Programmes can be met through (a) participation in the Annual National Report on Schooling in Australia, (b) the provision of agreed reports, and (c) participation in evaluations of the outcomes of programmes of financial assistance. The Act allows for funds to be withheld from jurisdictions if the Minister is not satisfied that accountability requirements have not been met.

Since the Act requires education authorities to account for NALSAS funds, it is worth examining some of its provisions in more depth, as they may provide some insights into how an outcomes framework for NALSAS might be implemented in future. For 2000, accountability for NALSAS was able to be satisfied by participation in the ANR and programme evaluations requested by the Commonwealth. The assumption made in this paper is that the NALSAS outcomes framework will continue to be reported through these mechanisms.

Sections 17-29 of Appendix E of the Act indicate that participation in the ANR comprises reporting against items set out in the Agreed Information Framework and participation in the sample studies endorsed by MCEETYA as part of the Framework. The framework reflects the National Goals for Schooling in the Twenty-first Century, and includes for 2000 a requirement for education authorities to report performance against the agreed literacy and numeracy benchmarks for Years 3 and 5. Since the National Goals, and therefore the Information Framework also, promotes student learning in all Key Learning Areas including Languages Other Than English (LOTE) (which includes the NALSAS languages), and Studies of Society and Environment (which covers studies of Asia), it may be reasonable to expect, as resources and circumstances permit over time, appropriate reporting across all Key Learning Areas. LOTE, including Asian languages, have yet to be included in the Agreed Information Framework in terms of student outcomes, reflecting greater priorities and the pragmatics of current data availability. A NALSAS outcomes reporting framework would provide an opportunity to advance the reputation of LOTE (and Asian languages in particular) as being accountable and concerned with improvement in programme delivery and student learning.

It is fair to say that while these legislative and administrative requirements provide a strong basis for an accountability framework and legitimacy for a national perspective in this area, the reality of the present reporting efforts falls short of what is desired. State and territory officials would agree that it is difficult for them even to provide participation data in a comparable form, let alone any data that would indicate the extent to which learning has taken place. It is relatively easy for jurisdictions to demonstrate financial accountability (in terms of compliance with contractual requirements), but much more difficult for them to demonstrate educational accountability. In its fullest expression, educational accountability requires information not only about what has been achieved, but why things are this way, and what will be done as a consequence. To do so may require considerable effort on the
part of the various jurisdictions (and reallocation of resources) to establish the appropriate systems and processes to collect, organise and present this information.

MCEETYA has established a National Educational Performance Monitoring Taskforce (NEPMT) to advance the measurement and reporting agenda in relation to the National Goals. The development of an outcomes framework for NALSAS needs to be compatible with the work of the ANR Taskforce and the NEPMT. While NEPMT has not currently scheduled work in the LOTE area as a priority, leadership in this area by NALSAS could be beneficial for enhancing the credibility of this Key Learning Area as a whole.

Further impetus for the development of an outcomes framework comes from the impending Commonwealth funding sunset provisions for NALSAS (in 2002) and other aspects of LOTE (in 2004). Any arguments for renewal or continuation of NALSAS funding will need to be based firmly on the demonstrated successes from the programmes and activities initiated through past and current NALSAS funds. Even if NALSAS funding were to be withdrawn or reduced, States and Territories would need a more efficient means for tracking the successes of their efforts (in, for example, the long term effects of professional development funded through NALSAS, or the production of curriculum and syllabus documents) in terms of their impact on students, for their own internal reporting and management purposes.

3. The NALSAS Strategic Plan and reporting and accountability framework

As noted earlier, this paper derives from the strong focus on outcomes and accountability identified as one of four areas for development in the NALSAS Phase 2 Strategic Plan 1999-2002. This focus had two objectives:

- To develop an agreed framework for collecting and reporting on the outcomes of the NALSAS strategy; and
- To examine approaches for measuring levels of improvement of students’ skills and understanding of NALSAS languages.

Several strategies are proposed for achieving these objectives. The first proposes the development of a framework for the collection of data about student participation in Asian Language courses and other learning experiences. This framework, originally defined in the Appendix to the Strategy, and recently amended in the ANR Guidelines for 2000, is designed to capture the following data:

- Number of schools offering at least one NALSAS language by level of schooling
- Number of schools offering each NALSAS language by level of school
- Number of students studying each NALSAS language by year level, by hours of instruction per week (Primary)
- Number of students studying each NALSAS language by year level, by hours of instruction per week (Secondary),
- Number of students studying each NALSAS language in Years 11 and 12, by gender; and
- Number of students studying each NALSAS language by mode of delivery.
This data is important in informing understanding of the reach and uptake of the NALSAS languages, and for quantifying the extent to which the original strategic targets have been met. It will be reported for the first time in the 2000 ANR. To be able to report these data nationally will be a considerable advance on the current situation.

This information will be helpful from a management perspective for monitoring the implementation of the Strategy. However, it is limited in its ability to provide a true evaluative picture of the status of student accomplishments in terms of Asian language acquisition. It does not provide any significant evidence of student’s ability to communicate in Asian languages, or their appreciation of cultural aspects of languages and Asian societies, for example. Nor is it able to describe the unique contributions of schools or systems over other influences on students’ knowledge and understandings, such as the media, community associations or native speaker experience. In other words, it does little to indicate the value added by schools for the dollars being spent.

Both the NALSAS Strategic Plan, and the Commonwealth guidelines provide little detail about their expectations for student language learning, implicitly or explicitly. It is impossible to develop an adequate framework for reporting student learning outcomes without grappling with the issue of appropriate standards, whether we call them descriptors of performance, benchmarks or any other term. Standards or expectations for performance, whether implicit or explicit, always underpin assessment tasks and judgment rubrics that provide the data which is reported as “outcomes”.

The report *Advancing Australia’s Languages* notes that some work has been funded under the NALSAS strategy to define minimum proficiency outcomes, but it is unclear how systems propose to implement these standards in schools, or how they relate to systemic syllabus and curriculum documents. Issues in relation to use of proficiency outcomes statements are discussed in more detail in a later section of this paper.

The COAG working group in 1995 recognised that it was pointless to develop an expanded school-based program in Asian languages unless there was some means of measuring its long-term effectiveness. The quantitative targets for participation above provide one basis of measuring program outcomes, but these must be integrated with appropriate qualitative measures capable of assessing proficiency outcomes. The working group believed that none of the options for assessment available at the time (1994) were appropriate for a school-based program. They recommended urgent work in developing agreed proficiency scales, testing and reporting mechanisms for the four priority languages.

The concept of proficiency scales, how they might be defined – and how they might be written – has been hotly debated over the past few years (see for example, Scarino, 1997). Consideration of the options for an outcomes reporting framework discussed below provide an opportunity for continuing the debate about how language proficiency should be assessed and reported.
The COAG report recommends a proficiency scale based on what might be called “real-world” competency, as follows:

- Survival proficiency
- Minimum social proficiency
- Minimum vocational proficiency
- Useful vocational proficiency.

Other proficiency scales are clearly possible. Commonly used scales are those that link what we believe students should know and be able to do at various grade levels, age levels or stages of schooling. Other scales based on some notion of standards that rank students from novice to expert in some way are also used internationally. Developing a national proficiency scale in Australia will be complicated by the fact that different States and Territories use different organising frameworks in their syllabuses and curriculum documents.

4. The place of Asian Studies and Languages in State/Territory curriculum frameworks

A complete curriculum mapping exercise is beyond the scope of this paper. However, it is worth noting briefly the current opportunities for Asian language learning in the various States and Territories.

Basic notions of assessment validity hold that tests should measure what students have been taught. In the best-case scenario, this means assessment should be based on the curriculum received by students, or at least that which has been delivered to them. In practice, it is more usual that assessment tasks should be based on the curriculum that students have at least theoretically have had the opportunity to learn.

Table 1: What provision within the state’s curriculum is there for students to undertake study of four languages.

<table>
<thead>
<tr>
<th>State</th>
<th>Year 6/7</th>
<th>Year 10</th>
<th>Year 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>New South Wales</td>
<td>Syllabuses have been developed by the Board of Studies in all four Asian languages. However, the choice of language is made at individual school level.</td>
<td>Students must complete 100 hours of language study between Years 7-10, however, Asian languages are not compulsory. There are syllabuses in each Asian language area.</td>
<td>Students may select an Asian language as an optional subject for the HSC.</td>
</tr>
<tr>
<td>Victoria</td>
<td>LOTE is compulsory for all students from P-10 however schools determine which particular language/s are offered.</td>
<td>As for 6/7</td>
<td>Victoria has a target of 25% of students studying LOTE at Years 11-12 level. Currently approximately 19% of students study another language.</td>
</tr>
<tr>
<td>Tasmania</td>
<td>Grades 3-6 have the opportunity to join the funded LOTE Primary Implementation</td>
<td>Participation statistics maintained on LOTE website: <a href="http://www.discover.tased.edu">www.discover.tased.edu</a></td>
<td>At least one NALSAS language is offered at each of the Senior Secondary colleges.</td>
</tr>
<tr>
<td>State</td>
<td>Year 6/7</td>
<td>Year 10</td>
<td>Year 12</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>Western Australia</td>
<td>Policy of gradual implementation of compulsory LOTE. By 2000 all students in years 3–7 and by 2003 all students in years 8 to 10 will be studying one of the Department’s 12 priority languages, four of which are the NALSAS targeted languages. In 2000 all but 5 schools primary schools had implemented LOTE. Participation data reflects 51% undertaking Indonesian, Chinese or Japanese, 49% undertaking other languages.</td>
<td>Compulsory implementation scheduled as follows: 2001 – targeting Yr 8 2002 – targeting Yr 9 2003 – targeting Yr 10 Currently 78 out of a total of 95 secondary schools (82%) offer a NALSAS targeted language.</td>
<td>Elective subject for Years 11 &amp; 12.</td>
</tr>
<tr>
<td>Northern Territory</td>
<td>Policy states that all students will have access to the study of another language as part of their general education, through regular school programs or after-hours classes. Time allocations across years not specified. Some schools have language programs running (and to varying extents), other schools don’t. At upper primary level in 2000 Indonesian was quite widely taught in urban schools, Japanese was taught to a limited extent, Chinese was available at one school. Korean is not taught in NT schools.</td>
<td>As for 6/7. Generally an elective subject at Yr 10. At Yr 10, in 2000, Indonesian and Japanese were taught in a number of schools, Chinese was not taught. As an elective subject it is likely that some schools may not have run Yr 10 classes due to minimal enrolments.</td>
<td>Elective subject for Years 11 &amp; 12. In 2000 Indonesian and Japanese were taught at Yr 12 in a number of schools, Chinese in two schools.</td>
</tr>
<tr>
<td>Queensland</td>
<td>Compulsory participation is required in Years 6,7. Opportunities are available for all schools to take up the four NALSAS languages.</td>
<td>Compulsory participation is required in Year 8.</td>
<td>Elective subject for Years 11 &amp; 12.</td>
</tr>
<tr>
<td>State</td>
<td>Year 6/7</td>
<td>Year 10</td>
<td>Year 12</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Australian Capital Territory</td>
<td>The ACT operates a system of school based curriculum development. Curriculum documents are based on the ACT LOTE Curriculum Frameworks. Courses are designed along communicative lines set out in the LOTE Curriculum Framework incorporating across curriculum perspectives, key competencies and work related needs.</td>
<td>As for 6/7.</td>
<td>Courses are developed in schools under relevant subject area curriculum documents called Course Frameworks and are accredited by the ACT Board of Senior Secondary Studies. The school based approach to course development ensures that teachers have a primary role in developing courses that reflect their teaching styles and are directly relevant to their students’ context.</td>
</tr>
<tr>
<td>South Australia</td>
<td>In South Australia there is a mandated curriculum framework, the recently developed South Australian Curriculum Standards and Accountability (SACSA) Framework. The Framework covers birth to year 12 and includes languages as one of the eight learning areas. The SACSA Framework articulates the scope and standards for languages from Reception to year 12. The R-12 scope and R-10 standards for the languages learning is described through three strands: * communication * understanding language * understanding culture The scope and standards for the languages learning area will be articulated according to three groups of languages: -Alphabetic (this includes Indonesian) - Non-Alphabetic</td>
<td>As for 6/7.</td>
<td>As for 6/7. Additionally, Year 12 standards are described in terms of Essential Learnings. The specific learning area requirements for senior secondary languages is articulated through the Collaborative Curriculum and Assessment Framework for Languages as part of the South Australian Certificate of Education, under the jurisdiction of SSABSA, the Senior Secondary Assessment Board of South Australia.</td>
</tr>
<tr>
<td>State</td>
<td>Year 6/7</td>
<td>Year 10</td>
<td>Year 12</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td></td>
<td>(includes Chinese and Japanese)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-Australian Indigenous</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Both the Non-Alphabetic and Australian Indigenous are currently under</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>development with anticipated completion by September 2001. Studies of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asia is a priority as a cross-curriculum perspective as part of the</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SACSA Framework.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victoria</td>
<td>LOTE is compulsory for all students from P-10 however schools determine</td>
<td>As for 6/7</td>
<td>Elective subject for Years</td>
</tr>
<tr>
<td></td>
<td>which particular language/s are offered.</td>
<td></td>
<td>11 and 12.</td>
</tr>
<tr>
<td>Tasmania</td>
<td>Asian Languages study is possible at individual school level. Syllabus</td>
<td>Non-compulsory subject for Years 7-10.</td>
<td>Elective subject for Years</td>
</tr>
<tr>
<td></td>
<td>is based on the Statements and Profiles</td>
<td></td>
<td>11 &amp; 12. Board developed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>syllabus</td>
</tr>
<tr>
<td>Western Australia</td>
<td>Policy of gradual implementation of compulsory LOTE. 2000 - all primary</td>
<td>Compulsory implementation scheduled as</td>
<td>Elective subject for Years</td>
</tr>
<tr>
<td></td>
<td>schools Years 3 to 7 implemented LOTE with exception of 5 remote</td>
<td>follows: 2001 – targeting Yr 8</td>
<td>11 &amp; 12.</td>
</tr>
<tr>
<td></td>
<td>aboriginal schools. Participation data reflects 51% undertaking</td>
<td>2002 – targeting Yr 9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Indonesian, Chinese or Japanese, 49% undertaking other languages.</td>
<td>2003 – targeting Yr 10</td>
<td></td>
</tr>
<tr>
<td>Northern Territory</td>
<td>Policy indicates that Languages is a Key Learning Area. Policy has</td>
<td>As for 6/7</td>
<td>Elective subject for Years</td>
</tr>
<tr>
<td></td>
<td>been interpreted differently by each school and as a result some</td>
<td></td>
<td>11 &amp; 12.</td>
</tr>
<tr>
<td></td>
<td>schools include LOTE and some don’t.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Queensland</td>
<td>Compulsory participation is required in Years 6,7. Opportunities are</td>
<td>Compulsory participation is required</td>
<td>Elective subject for Years</td>
</tr>
<tr>
<td></td>
<td>available for all schools to take up the four NALSAS languages. Many</td>
<td>in Year 8.</td>
<td>11 &amp; 12.</td>
</tr>
<tr>
<td></td>
<td>have changed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian Capital</td>
<td>Schools determine</td>
<td>As for 6/7</td>
<td>Elective subject for</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>Year 6/7</td>
<td>Year 10</td>
<td>Year 12</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Territory</td>
<td>what languages will be taught.</td>
<td>Years 11 &amp; 12.</td>
<td></td>
</tr>
<tr>
<td>South Australia</td>
<td>In South Australia there is a mandated curriculum framework, the recently developed South Australian Curriculum Standards and Accountability (SACSA) Framework. The Framework covers birth to year 12 and includes languages as one of the eight learning areas. The SACSA Framework articulates the scope and standards for languages from Reception to year 12. The scope for the languages learning is described through three strands: * communication * understanding language * understanding culture The scope and standards for the languages learning area will be articulated according to three groups of languages: - Alphabetic (this includes Indonesian) - This is currently published and to be on the web by February - Non-Alphabetic (includes Chinese and Japanese) - Australian Indigenous Both the Non-Alphabetic and Australian Indigenous are currently under development with anticipated completion by September 2001. Studies of Asia is a priority as an across the curriculum perspectives as part of the SACSA Framework.</td>
<td>As for 6/7. Additionally, Year 12 standards are described in terms of Essential Learnings. The specific learning area requirements for senior secondary languages is articulated through the Collaborative Curriculum and Assessment Framework for Languages as part of the South Australian Certificate of Education, under the jurisdiction of SSABSA, the Senior Secondary Assessment Board of South Australia.</td>
<td></td>
</tr>
</tbody>
</table>
As the 1998 Evaluation of the Commonwealth School Language Program found, the original targets set for the proportion of Australian students studying Asian languages were not being met, nor do more recent figures suggest that much further ground has been gained (see Table 2). However, students in all States and Territories do have at least the potential opportunity to study Asian languages, as every state has developed or adopted syllabuses or curriculum frameworks that make studies of Asia and Asian languages possible. Whether or not a school makes these opportunities available to students depends on their local circumstances. As Table 3 shows, study of NALSAS languages is not mandatory in any jurisdiction.

**Table 2: Participation in NALSAS languages**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>25110</td>
<td>59563</td>
<td>108848</td>
</tr>
<tr>
<td>Chinese</td>
<td>8343</td>
<td>14066</td>
<td>19970</td>
</tr>
<tr>
<td>Indonesian</td>
<td>5824</td>
<td>30903</td>
<td>73142</td>
</tr>
<tr>
<td>Korean</td>
<td>0</td>
<td>228</td>
<td>416</td>
</tr>
<tr>
<td><strong>Year 12</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>3115</td>
<td>5451</td>
<td>5381</td>
</tr>
<tr>
<td>Chinese</td>
<td>2041</td>
<td>2952</td>
<td>2361</td>
</tr>
<tr>
<td>Indonesian</td>
<td>1186</td>
<td>1546</td>
<td>1869</td>
</tr>
<tr>
<td>Korean</td>
<td>0</td>
<td>247</td>
<td>248</td>
</tr>
</tbody>
</table>

Source: Advancing Australia’s Languages Overview Report, 1998, pp. 28-29. Primary data based on schools in Victoria, Tasmania and Queensland only.

**Table 3: Is study of an Asian language mandatory in this State?**

<table>
<thead>
<tr>
<th>State</th>
<th>New South Wales</th>
<th>Victoria</th>
<th>Tasmania</th>
<th>Western Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not compulsory</td>
<td>Students must complete 100 hours of language study between Yrs 7-10, but can be any language.</td>
<td>Not compulsory</td>
<td>LOTE studies are compulsory, Schools chose which language to implement from a list of 12 priority languages, 6 of which are Asian languages.</td>
</tr>
<tr>
<td><strong>Victoria</strong></td>
<td>While the study of LOTE is mandatory, the study of an Asian language is not. An Asian LOTE is provided in 62% of government schools and approximately 53% of students learning language learn an Asian language.</td>
<td>As for 6/7</td>
<td>Victoria has a target of 25% of students studying LOTE at Years 11-12 level. Currently approximately 19% of students study another language.</td>
<td></td>
</tr>
<tr>
<td><strong>Tasmania</strong></td>
<td>Not compulsory</td>
<td>Not compulsory</td>
<td>Not compulsory</td>
<td></td>
</tr>
<tr>
<td><strong>Western Australia</strong></td>
<td>LOTE studies are compulsory, Schools chose which language to implement from a list of 12 priority languages, 6 of which are Asian languages.</td>
<td>Compulsory implementation scheduled from 2001 to 2003. Schools chose which language to implement from a list of 12 priority</td>
<td>Elective subject for Years 11 &amp; 12.</td>
<td></td>
</tr>
</tbody>
</table>
Currently 51% study either Indonesian, Chinese or Japanese. Languages, 6 of which are Asian languages. Currently 82% of secondary schools offer an Asian language.

| Northern Territory | Policy states that all students will have access to the study of another language as part of their general education, through regular school programs or after-hours classes. Schools are encouraged to choose from those identified for mainstream education, namely Indigenous languages, Chinese, French, German, Indonesian, Italian, Japanese, Greek. | As for 6/7 | Not compulsory |

| Queensland | Language study mandatory in Yrs 6/7/8, but not specifically Asian languages | Not compulsory | Not compulsory |

| Australian Capital Territory | Study a LOTE for all Year 7 students is compulsory except for one high school where it is an elective. Study of an Asian language is not mandatory. | Study of an Asian language is not mandatory. | Optional for Years 11 & 12 Study of an Asian language is not mandatory. |

| South Australia | The Languages other than English Plan 2000-2007 sets out future directions for the teaching and learning of languages. The goal of the plan is that by 2007 R-10 students in government schools will be learning at least one language other than English in quality programs that are an integral and valued part of a broad and balanced curriculum. Chinese, Indonesian and Japanese are identified as three of a range of languages to be supported. | As for 6/7 |
5. Current status of outcomes measurement of Asian Languages in States and Territories

In 1998, the evaluation of the Commonwealth School Languages Programme, *Advancing Australia’s Languages* noted a lack of mechanisms to measure improved proficiency in language learning. The evaluation found that:

“There is an inherent contradiction in the current NALSAS Year 12 targets which appear to set ‘proficiency’ goals for the Strategy, and NALSAS’s implementation scope which some systems vigorously argue is only concerned with Years 3 to 10. There is an implicit belief that systems can not be held responsible for enrolment trends in post-compulsory years nor the improved proficiency associated with further study. The lack of proficiency measures also leads systems to report on increases in terms of enrolments.

In primary schools where most enrolment growth is occurring, assessment in all subjects is conducted by the class teacher, with little if any moderation or central reporting of ‘proficiency’. In addition, Independent schools, which make a significant contribution to enrolments in some languages, are under no obligation to report to a ‘system’.

At secondary school, other than Year 12 results and the National Australia Bank / ACER language certificates available in some languages at lower secondary school, no general measures of proficiency are being used.”

It appears that little has changed in 2001. In preparing this paper, a brief survey of LOTE and NALSAS managers in each state and territory was conducted. This survey suggested that few systemic measures of language learning are conducted at present and that even fewer are publicly reported. The following table summarises the findings of the 2001 survey.

Table 4: Systemic assessment and reporting of Asian languages

<table>
<thead>
<tr>
<th>State</th>
<th>Year 6/7</th>
<th>Year 10</th>
<th>Year 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>New South Wales</td>
<td>No assessments of student outcomes except at school level. A small-scale program evaluation, conducted as an “in-house” exercise is being conducted which includes an attempt to compare student performance in schools with targeted programs and control schools and over time.</td>
<td>No public reporting of outcomes, although students choosing Asian languages are given a grade (A-E) for the subject for their School Certificate. This is determined at the school level. Teachers are expected to assess students against the standards implicit in the course descriptors (see appendix for an example) but teachers interpret the standards in different ways. Even if it were possible to aggregate these school-awarded grades to the state level, the result would most likely resemble a normal distribution.</td>
<td>Students taking Asian languages are assessed as part of the HSC. Because the exam has been norm referenced (until 2001), without access to the raw assessment marks and significant statistical manipulation, aggregate results at systemic level are meaningless. From 2001 the HSC will move to a standards referenced model, which in theory should mean that reporting of the percentage of students reaching particular proficiency levels should be possible. A mark scale is awarded based on performance relative to the norm population.</td>
</tr>
</tbody>
</table>
distribution. The planned development of Yrs 7-10 syllabuses with outcomes for the NALSAS languages will give an enhanced capacity to measure student performance.

| Victoria | The CSF11 provides levels and outcomes at 6 levels, including extension at Level 6, for students in Years P-10 in the 8 key learning areas including LOTE. Schools are provided with guidelines for assessment and reporting in all curriculum areas, however, schools are responsible for implementing their own assessment and reporting practices. Teacher Support Materials (TSM) for LOTE, designed to assist with implementation of CSF11, provide examples of assessment tasks and strategies. TSM are currently in development for Chinese, Japanese and Indonesian. | As for 6/7 | At VCE level (Years 11 and 12) student performance in all languages, including the NALSAS languages is assessed in accordance with the VCE and managed by the Board of Studies (BOS). |
| Tasmanina | Moderated teacher assessments against syllabus outcomes are collected systemically from schools in the program (Grades 3-6). Reporting is in terms of percentage of students meeting the criteria for each level, according to the National Profiles. This data is reported to DETYA as part of the accountability report, but is not otherwise published. | Moderated teacher assessments against syllabus outcomes and standards are collected from all schools. Results are monitored by TASSAB (Tasmanian Secondary Assessment Board.) The board maintains records of student achievement and issues statements of qualification. | Year 12 students studying languages are examined as part of the HSC by TASSAB (Tasmanian Secondary Assessment Board) The board maintains records of student achievement and issues statements of qualification |
| Western Australia | By 2004 all schools will have implemented the Curriculum Council’s Curriculum Framework and students will be achieving to some | As for Year 6/7 | Year 12 students studying languages are examined as part of the WACE. |
degree the LOTE outcomes in this Framework. Teachers design learning programs to enable all students to achieve the outcomes. The Education Department’s Outcomes and Standards Framework is the progress map for measuring students’ achievement of the outcomes. This map describes 8 levels of achievement. By 2005 schools will be reporting student progress in LOTE against the 8 levels. Standards and Effectiveness Branch developed standardized tests for Indonesian at Year 7 level in 1997 and Japanese at Years 7 and 10 in 2000. Japanese to be tested statewide in 2001.

<table>
<thead>
<tr>
<th>Northern Territory</th>
<th>No system-level assessment and reporting of language learning outcomes. (Information on student participation ie enrolments is collected as part of the Department’s annual Languages Census.)</th>
<th>As for Year 6/7</th>
<th>Senior secondary accredited courses involve specific assessment procedures. Results are reported on students’ NT Certificate of Education.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queensland</td>
<td>Assessment is undertaken by individual teachers, however, guided by outcomes defined within the syllabus. This data is not aggregated or reported publicly. Participation data only is collected.</td>
<td>Opportunities are available for all schools to take up the four languages. Many have converted. Assessment is undertaken by individual teachers, however, guided by outcomes defined within the syllabus. Participation data only is collected. This data is not reported.</td>
<td>Years 11 &amp; 12 assessments are moderated across the State via the BSSS.</td>
</tr>
<tr>
<td>Australian Capital Territory</td>
<td>Student outcomes are assessed and reported on to the Department in terms of outcomes from the LOTE Profile and Frameworks.</td>
<td>As for Year 6/7</td>
<td>The Course Frameworks are system wide documents that provide the essential basis for course development and</td>
</tr>
</tbody>
</table>
It is both fortunate and unfortunate that there is no strong tradition of national outcomes reporting in the Asian languages area. It is unfortunate, in that there are no existing baselines against which progress can be measured. It is fortunate, in that the absence of existing measures means that we do not have to be constrained by existing thinking or testing regimes that are difficult to change. The following section considers some possible options for collecting national student outcomes data in this area.

6. Some preliminary considerations in planning an outcomes reporting framework

In thinking about what outcomes might be reported and how they might be reported, it is essential that three important questions be answered:
1. What do we (ie, educational policy and decision-makers) most need to know about student performance in NALSAS programmes?
2. How will the information be used?
3. What is the cost (in financial and non-financial terms) of collecting this information, and what is the cost of not having the information?

The answers to these questions have significant implications for the design of the outcomes reporting framework and the data gathering activities that support it. Some forms of outcomes data gathering (ie assessment instruments) will be incompatible with some information objectives. For example, normatively-based assessments, whose principal function is to rank students, do not give information about distributions of performance. It is possible that the NALSAS Taskforce might want to know, for example:

- the proportion of students who have mastered certain skills, understand certain concepts or who have acquired knowledge of particular subject matter
- in which language(s) students perform best in these skills and in which languages they perform less well
- whether student performance, relative to some standard, is increasing or decreasing over time
- whether student attitudes towards language learning change over time and what factors appear to influence these attitudes
- whether particular forms of organisation of learning (such as immersion classes) appears to be associated with higher or lower levels of student learning outcomes
- whether the kind of initial training/professional development of teachers is associated with higher or lower levels of student learning outcomes
- whether the availability of resources of various kinds (such as access to satellite broadcasts) are associated with higher or lower levels of student learning outcomes
- whether certain school policies, level of community involvement, and systemic support are associated with higher or lower levels of student learning outcomes
- whether the proportion of native speakers of the language in the class/school community has an influence on the language acquisition of non-native speakers or the class as a whole.

These are just a few examples of the kinds of questions that could be asked. From these examples though, it can be seen that an outcomes reporting framework can do much more than simply report that say, X % of students have reached a particular (necessarily arbitrary) benchmark or standard.

Performance measurement systems have multiple goals, and multiple audiences. At the most basic level, they provide an aggregation of “facts” about the status of a project, programme or strategy. These facts might help to describe, or give an account of the functioning of the thing of interest. Reporting the NALSAS outcomes in the ANR is an example of a communicative view of accountability, in terms of “giving an account” of the programme. This communication function is important. There are many stakeholders in Asian language education, each of whom has different information needs. The organisation, accessibility and clarity of presentation of data are important determinants of the level of use of performance data.
While it is essential that the communication function is properly developed, performance measurement systems are most often thought to have a stronger accountability role, in allowing demonstration that:

- contractual obligations have been met
- funds have been acquitted properly
- programs have been implemented as intended
- programs are achieving their intended outcomes or targets have been met.

Public accountability demands such measurement to monitor compliance, but program managers also acknowledge that performance information is essential to guide improvement. According to Mitchell, (1989) there are four motivations for measuring educational phenomena: (a) the pragmatic – does it work? (b) the moral – is it good? (c) the conservative – is it necessary? and (d) the rational – does it make things better? Language education involves a complex, evolving set of models, practices, policies and activities that overlap and interact over time. Despite the fact that we have strong emerging evidence of “what works” in terms of second language teaching, few would claim that any program or strategy has yet reached its ultimate stage of development or produces foolproof results.

Stakeholders with different roles in education systems have different needs for information. The same questions, and therefore the same data sources, are not necessarily appropriate for providing answers to these information needs. The information needs are not necessarily hierarchical, nor are they additive. Conceptually, these needs can be grouped into three aspects, as follows:

- Policy information
- Management information
- Strategic information.

The NALSAS Taskforce will need to consider which of these types of information is most needed from the outcomes framework. Given the likely cost of obtaining national outcomes data, it would seem appropriate to also consider developing a performance information system that would inform the development of the next possible iteration of the NALSAS Strategic Plan. Students outcomes testing can provide performance data, such as the percentage of students who have demonstrated a particular outcome, say the ability to ask directions. However, the development of an outcomes reporting framework presents an opportunity to go beyond this.

A performance information system is concerned not only with collecting the data, but comparing, presenting, reporting and disseminating the data in such a way as to invite reflection, evaluation and action. Gathering project input or even outcome data is not sufficient if nothing is done with it apart from being filed. Again, to be a system, reporting must be planned, appropriate to its intended audience, and have internal coherence and consistency.

Performance management systems take the information and use it in a deliberate way to support decision making. On the basis that “What gets measured, gets done”, information can be used to clarify and communicate goals, focus on areas in need of attention, reward and sanction performance levels outside certain criteria, and so on.
Performance measurement can be connected to action if it is part of a monitoring and management information system. The important features here are feedback loops – direct links between the information provided in the monitoring system and decisions taken at program or system level to effect changes. This allows management of the program or system performance.

However, not everything within a program or system needs to be measured and monitored to provide performance information. The costs and time consumed make this impractical. The core task in developing an effective performance monitoring and management information system is the identification of a set of key performance indicators – key measurements that reflect system performance.

The NASLAS Taskforce has the opportunity to consider what kind of performance management system might be developed at the same time as the outcomes framework to assist them with monitoring the implementation of the Strategy. It is possible to design information gathering processes that can provide information at multiple levels. It is much easier to design these processes before designing the data gathering mechanisms than it is to cobble them together from whatever is available at a later stage.

7. Possible means for assessing and reporting student outcomes

There are only a limited number of ways in which student outcome data in relation to NALSAS objectives can be collected and reported on a national basis. Two approaches are possible – either to use existing data, or to collect new data. The options discussed below are all subsets of these two approaches. Each of these options has advantages and disadvantages in terms of their feasibility and desirability from a conceptual and pragmatic point of view.

Option 1

**Report systemic data collected as a by-product of internal administrative arrangements or systemic outcomes measurement regimes** (which might include public exams, “basic skills tests” or aggregated school-based assessments) on a state by state basis.

In this option there would be no attempt to obtain comparability or consistency between jurisdictions. Data would be reported “as is”, in whatever form could be supplied by the states and territories, in a manner similar to that in which student outcomes have been reported in the Annual National Report (ANR) in the past. There would be no attempt to create synthetic tables, such as are now shown for literacy outcomes in the supplement to the 1999 Annual National Report on Schooling. As Table 3 indicates, there are likely to be gaps for particular languages and levels in various jurisdictions.

The major advantage of this option is that it is relatively simple to conduct, requiring only an annual survey of jurisdictions (provided that “gaps” in the data are expected and acceptable). It is relatively unobtrusive on systems and schools, and relatively inexpensive since it does not require the generation of
“new” data. This survey could be conducted annually, although as with all such data collections, depending on timing of the survey, it may be reflecting the previous year’s performance rather than the current year. In the absence of better data, it at least provides a starting point from which the data collection deficiencies become obvious, and data needs and issues can be clarified.

The major disadvantages of this option are that, at present and for the foreseeable future, few jurisdictions systematically collect student outcomes data that could be used in this way. Even where outcomes data is collected, it is often unintelligible at an aggregate level, since the assessment model is primarily intended for certification of individuals. For example, up to now it has not been possible to report trends over time or to make judgements about the “quality” of students’ results from the NSW Higher School Certificate examination, because the scaling procedures used result in all courses having the same “average” and distributional properties. As the exam was normatively based within the year cohort, it is not possible at the macro level to report how students have performed against certain criteria or benchmarks. (NB from 2001 the NSW HSC will be based on a standards referenced framework, which may mean that it will be possible to make inferences on student performance at an aggregate). However, the fact remains that it would be difficult to obtain comparable data for all States and Territories.

**Option 2**

**Report systemic data collected as a by-product of internal administrative arrangements or systemic outcomes measurement regimes on a state by state basis, equated on a common scale.**

The second option is similar to Option 1, except that data collected by jurisdictions is equated in some way to obtain “scores” which can be compared between languages, levels and states and territories. This concept has been explored thoroughly in the psychometric literature and underpins many attempts to develop nationally comparable data. Using this process, it is possible, for example to place the NSW Basic Skills tests in literacy, and the Victorian CRTs on a common scale.

The advantage of this option over Option 1 is that direct comparisons are now possible, and it is easier to explain student performance in terms of well-defined standards. The disadvantages are that there is little data available which could be used in this way, there is little agreement as yet among jurisdictions about what the common standards might be, and the methodology of equating and calibrating different assessment is time-consuming and therefore expensive to develop.

In the absence of any readily available or useable existing data, it is likely that new data will need to be collected. There are several options here.

**Option 3**
Use the ACER Language Proficiency examinations as the basis of a national test.

This option has the advantage that the tests have already been developed, and have known properties. There is an established administrative process for the conduct and administration of the assessment. Because the assessment is primarily multiple choice, the scoring and analysis process can be conducted quickly and cheaply. The process can be completed without placing unnecessary demands on the classroom teacher. The assessment will provide objective and consistent data across the country. Because assessment procedure is designed for the award of a certificate, there is an additional incentive for schools and students to participate.

The disadvantages are that the content of the examinations are not necessarily aligned to the curriculum in any jurisdiction. There are concerns about the validity of the assessment, both in terms of the content of the test and how it is tested. There is no oral component to the test, for example, yet this is a major component of the junior curriculum. The assessment measures only a limited range of the curriculum objectives. There are always concerns for the validity of conclusions drawn when assessments designed for certifying individuals are applied as measures of systemic performance. Most particularly, since the test can be regarded as “high stakes”, in that it results in the award of a certificate, teachers may “teach to the test”, and thereby perhaps distort the intentions of the curriculum. The current range of examinations also do not relate to all stages of schooling.

Option 4

Develop a national standards framework against which teachers assess their students. The framework would have implicit a proficiency scale, with descriptors of performance at each level. Teachers would be free to use whatever evidence they wished to make the judgement about whether a particular student had been able to demonstrate mastery at a particular level. The framework implicitly accepts the notion of the progression of learning, that language acquisition is cumulative, that there are multiple entry points to language learning, and that students can demonstrate mastery of particular objectives in different ways. The framework could be written in either generic terms or be language specific. The option recognises the importance of teachers’ professional judgment in assessing complex concepts.

The major advantage of this option is that administratively, collection and reporting of the data is simple. The assessment regime could be broken down to provide scores in a number of strands, such as Listening, Speaking, Reading and Writing. All the class teacher would have to do is complete a simple pro-forma, such as the following.

<table>
<thead>
<tr>
<th>Writing</th>
<th>Reading</th>
<th>Speaking</th>
<th>Listening</th>
<th>Socio-Cultural</th>
</tr>
</thead>
</table>

Please indicate the number of students in your class who have reached each level in each strand.
Teachers would be given exemplars of what students are expected to know and be able to do at each stage to assist them make this judgment. These exemplars could be more or less explicit, including being language specific.

This option fits well with teacher’s beliefs and current trends in classroom assessment. It would reinforce teacher’s understanding of standards in language learning. The major costs of this option come in the development of the standards framework, which, once developed, could be used in subsequent years. Teacher time costs will be absorbed into current school operating costs.

The biggest disadvantage is that developing a national standards framework is far easier said than done. Because states and territories differ widely in what they believe are appropriate standards and content for each of the languages, (ie what the descriptors or indicators should be at each level), the development of an agreed framework will be time consuming and frustrating, if past efforts to do so are a guide. For example, the NALSAS funded descriptors of exit outcomes, or indeed the National Statements and Profiles have not been adopted or implemented in the majority of jurisdictions.

Another disadvantage is that while this option can produce rich data at the school level, it provides scant detail at systemic or national level. However, the greatest disadvantage of this method is that its level of reliability is suspect. Despite being provided with guidelines and aides-memoire, teacher judgments are individualistic, and are based upon their internalised system of standards, drawn from their own accumulated knowledge and experiences. There is some evidence that teachers are not good at designing assessment activities that tap higher-order thinking, and allow prejudices about individuals and the other factors to enter into the judgment process. These factors can be reduced by using moderation processes (such as used in the assessment process in Tasmania) where panels of teachers make judgments on the basis of work samples and other evidence, which provides greater consistency in the judgments. However, while establishing moderation panels also has a good spin-off effect in providing a professional development activity for teachers, it also significantly adds to the cost of the exercise. Establishing nationally representative, cross-jurisdictional moderating panels will be logistically complex.

Option 5

The fifth option is to conduct a specially-designed national sample study. There are several precedents for this method in recent years, including studies
of students’ understanding of Information Technology, Civics, and knowledge and understanding of Asia.

Our view of a sample study is that it need not be limited to a simple paper and pencil test that produces a summary score as the means of reporting. A sample study could (but doesn’t necessarily have to) include some element of teacher assessment, and gathers student and school information simultaneously. Including some component of teacher participation in the process may help to increase the perception of the benefit of the study to schools and teachers, if it can help to develop their own knowledge and understanding of their students or their teaching practice. Other than at Year 12 level, many teachers have no external frame of reference to answer the question “are my students learning as much as they should”?

The advantage of a sample study is that a great deal of inferential information can be gathered from a relatively small sample of schools and students. Instruments and processes can be developed that have high levels of credibility among varied stakeholders. The methodology of sample studies is often easier to explain than equating studies. This can help promote greater acceptance. Further, sample studies can be relatively easily replicated in subsequent years to measure progress as well as status of learning, particularly if item banks rather than single assessment tasks are developed. Perhaps most importantly, for jurisdictions where there are currently no outcome measurement regimes in place, participation in a sample study can be a cheaper and more practical method of obtaining student outcomes data than developing their own, perhaps more comprehensive data systems.

Sample studies can be conducted at relatively low cost, depending on the size of the sample desired. Costs increase greatly if the survey is to be robust enough to provide disaggregations at jurisdiction level or for population sub-groups, such as native speakers, gender, and so on. Costs will be increased also if more than one language is tested.

The disadvantages of this option is that participation in external testing is often resented by schools, seen as intrusive and burdensome, and irrelevant to their real work. There is a long history of fear and apprehension by teachers, schools and education systems that information collected from testing programs will be used in a punitive manner, or reported in the press to their detriment. This is a problem not limited to Asian languages assessment, and can be overcome to some extent by providing detailed information to schools that assists them to further provide quality teaching and learning. More practically, there is a need to ensure that the assessment instrument is relevant to the curriculum of all jurisdictions (so that all students have an equal chance of demonstrating mastery). Similarly, there is a need to ensure that the assessment is capable of differentiating between students at different levels of proficiency, and is statistically valid, reliable and fair. The process of assessment development and approval can be lengthy. Rarely does the development cycle coincide with convenient or appropriate opportunities for implementation within the school year. There can be significant costs incurred in the development, administration and analysis of appropriate instruments.
Despite these disadvantages, Option 5 may be the most feasible method of proceeding (it may also be possible to combine some aspects of school-based assessment and sample testing). Options 1 and 2, while perhaps cheaper and simpler, are not viable given the absence of reliable or useable data in most jurisdictions. This is not to say that undertaking a sample study is easy. Strong leadership from the NALSAS Taskforce will be required to deliver the consensus among states on conceptual and practical issues necessary for the study to proceed. A sample study, which combines elements of student assessment with surveys of school, teacher and systemic factors provides the best opportunity for a useful evaluative mechanism (i.e., a performance information system) that can inform both accountability and program improvement. Well-designed information systems also facilitate communication about and celebration of the achievements of schools and systems in the focus area, as well as exposing areas in need of further development.

The following section gives further consideration to what might be involved in developing a sample study in this area.

7. Issues in developing Asian languages sample study instruments

7.1. Who should be tested?

The first issue to be resolved substantively in a sample study is who should be included in the sample. From a wider strategic point of view, there is some appeal to sampling the entire student population (whether or not they had been exposed to Asian language instruction) to obtain a measure of the “stock” of skills in Asian languages at a particular point in time and at any particular grade level. Such an assessment would be able to be used to demonstrate the gap between actual and desired levels of proficiency in the Australian student population. However, such an assessment would hardly be fair if we were concerned with knowing the effectiveness of language teaching programs, or the extent to which students had benefited from NALSAS-funded programmes. This latter purpose appears to be more closely aligned with the Commonwealth’s accountability requirements and the intentions in the NALSAS Strategic Plan.

If we accept that evaluation of effectiveness or knowledge and skill acquisition are the prime objectives of the sample study, then the question arises of how much instruction is necessary to validly reflect students’ opportunity to learn. We would expect, for example, that there would be significant differences in the levels of proficiency of those who had one year of study to those who had four years of continuous study. There may also be differences between those who have received the same overall amount of instruction (in terms of hours of instruction), but some have received say, an hour a week over three years, while others may have intensive lessons in a single year or semester. We would also expect differences between those who had no prior exposure to the language and those for whom this is their first or home language. Do we include those who have ever studied the language within a particular stage of schooling, or only those who are currently studying? This later point is particularly important in junior secondary school, for example in NSW, where it is most likely
that the majority of students receive language instruction in Years 7 and 8 (other than those who may select an Asian language as an optional area of study).

These variations are not a difficulty, provided that the population to be assessed is clearly defined. They can, in fact, provide an opportunity for comparing results obtained under different conditions. Specific rules for inclusion or exclusion from the sample can be written, and applied proactively at the school level or retrospectively at the analysis stage, or all students can be included with appropriate identifiers that allows later analysis. It could be decided, for example, to include within the sample only those who are currently studying the language and who have received at least two years or 100 hours of language teaching (or to ask background questions of students that allows the extent or type of instruction to be identified). The parameters can be as broad or restricted as desired.

It is usual in sample studies to take measurements at defined points in the schooling continuum, either at defined age levels or grade level, usually at the end of particular stages, (such as end of primary, junior secondary, senior secondary, which occur at different grade levels in different jurisdictions). The most sensible option would be to choose points that are the most inclusive, and where students have the best opportunities for demonstrating success. The end of schooling has some appeal in this regard because it gives the longest time for students to develop the required skills and understanding. It also reflects the emphasis on language learning as a cumulative experience, and fits with the concept of “exit outcomes” which underpin many jurisdictions’ curriculum documents.

Despite the desirability of having end of schooling data – which should represent the “best” levels of performance to be expected, since these students have both the longest exposure to the language and have chosen to specialise in them – it is often difficult to include Year 12 students in sample studies because of the reluctance of schools and school authorities to interfere with students’ preparation for final examinations. Students in Year 12 also typically finish formal lessons quite early in the school year. Late Year 11 can be substituted to provide an estimate. The same problems are also encountered, but to a lesser extent in Year 10.

At each year level, and with disaggregation no lower than state level, a total sample of under 4,000 students should be adequate for each language. A sample as low as 2,400 would be satisfactory if pure random sampling of students were possible. However, it is usually more convenient to use cluster sampling, since it is easier and cheaper to test whole or nearly intact classes. Even though this increases the number of students in the sample, it provides stronger school-level data.

Because of the substantial differences in the content of the curriculum and the methods of organisation of delivery of teaching and learning at primary and secondary level, assessment at two points might be most appropriate. In view of the work already completed to identify “exit outcomes”, assessment at both Year 6/7 and Year 10 may be most beneficial.

7.2. What should be tested?
Developing good assessment programmes for Asian languages is perhaps more complex than for more mainstream subjects. A fundamental principle underpinning the validity and reliability of assessments is that the assessment tasks should match closely the curriculum objectives in terms of both content and process. Determining just what the “curriculum” is, when a national perspective is adopted is the first hurdle to be overcome. States and territories differ in what is included in their syllabuses and curriculum statements at each level. It is necessary to first identify what elements may be common across the various jurisdictions, and whether there is sufficient commonality that any instruments developed from this core would be accepted as adequately reflecting what students have been taught in each jurisdiction.

A second fundamental principle guiding the development of assessment tasks is that if the curriculum is concerned with the development of skills, then the assessment should allow demonstration of those skills as directly as possible. An important emphasis in language learning is the development of oral and aural skills, particularly in the early years.

“Current practice, based on research into language acquisition, favours a communicative approach to the teaching and learning of authentic Japanese (i.e. learning of hiragana). The communicative approach focuses on the meaning and context of the communication before dealing with the structural (grammatical components of the discourse. It is considered most appropriate for situations primary students are likely to encounter.” (NSW Japanese K-6 Syllabus, p.6)

The implication here is that tests of language learning should address this oral language component. Paper and pencil tests may not be the best way to capture oral language proficiency. However, the alternatives to traditional tests (which include both interactive computer assisted testing, and one-to-one or small group oral testing by a trained test administrator) are both more expensive and time consuming.

Test content could include elements of:

- Values and Attitudes
- Written Communication – Reading
- Written Communication –Writing
- Spoken Communication – Listening
- Spoken Communication – Speaking

Spoken and written vocabulary (scripted and non-scripted languages), including character or word recognition and comprehension, language structures, (e.g. word order), and socio-cultural understanding and knowledge of the language as a system are all aspects which could be assessed in terms of their contribution to the ability to communicate effectively.

While a case can be made for the development of sample studies in each Asian language, cost factors are likely to be the major determinant of what is tested. Experts consider that the languages are sufficiently different from one another in terms of complexity and difficulty, and that the profile of learners in each language is sufficiently different, to warrant individual assessments. If priority has to be given to one language over another, a pragmatic way forward is to design the assessment for
the languages with the greatest number of students. At the present time, Japanese and Indonesian are the languages with the largest number of students. These languages also have the highest number of non-native speakers.

The construction of any assessment instrument is directed by two key questions:

- What knowledge, skills and understandings and values do we want students to develop, and
- What kinds of progress can we legitimately expect as students move through schooling?

The concept of measuring progress demands that a long-term view of learning needs to be adopted, particularly in the languages area, which is complex and problematic. Language learning is a cumulative process in which students need to retain for active recall and use all of the concepts they have previously learned (Scarino, 1997, p.8). If assessment instruments are to respect this notion and thus be more than a grab-bag of miscellaneous or fragmented chunks of knowledge. They need to be underpinned by a unified theory of learning. No single task can capture the multi-dimensionality of language learning.

We need to give a great deal of thought to the design of tasks which require higher order thinking, and the best configuration of tasks we can select to do justice to students learning, to allow them to demonstrate their best and to provide a basis for diagnosis.

We need to know more about the notion of performance itself and its relationship to the notion of knowing a language.

The problem in trying to define proficiency standards is that the curriculum (ie the intended outcomes) are often written in terms of “what”, but not “how well”. All language curriculum frameworks deal with aspects of communication in one form or other. For example, the Australian *Profile* level 3 statement is:

“Interacts in predictable social and learning situations, incorporating new language items into well rehearsed language patterns.

The Western Australian Learning Area Statement states that the major learning outcomes in languages are as follows:

1. *students comprehend and communicate in the target language through listening, responding and speaking*
2. *students view and read a variety of texts in the target language*
3. *students write a variety of texts in the target language*
4. *students develop an understanding of the culture(s) of the target language*
5. *students apply their knowledge of the structure of the target language to assist them to make meaning and create text*
6. *students use a range of communication skills and strategies to enhance their ability to convey and make meaning in the target language*

(Ministry of Education of Western Australia, u.p. p.3)
Both of these examples represent statements about what students do, with no attempt to describe the quality with which this is accomplished.

This is not an easy area for test developers to come to grips with. As noted above, it is also the area that teachers have the greatest difficulty in providing consistent judgments. What is considered adequate or exemplary demonstration of skills in oral language is particularly difficult, taking into account things such as inflections, fluency, pronunciation, dialects and so on. Students may be fluent, for example, in speaking in some contexts but not others. The task of the test developer is to include in the assessment a sufficient number of tasks that sample the domain so that we can be confident that the judgment about the level of performance demonstrated is correct, while at the same time balancing considerations such as the length of time available for testing, the likely circumstances in which the assessment will be conducted, and the impact of the test on the student.

The definition of appropriate test content is likely to be the area where there will be the greatest disagreement between stakeholders. There is no easy answer to this question. In writing the specifications for the sample study, the NALSAS Taskforce will have to seek consensus among stakeholders. There is inevitably a strong political aspect to assessment exercises. Assessment instruments send strong messages about what is valued within a subject area. Because there are limits to what can be included in any single assessment episode, priorities need to be assigned to some material over others.

7.3. When should students be tested?

The prime consideration in the timing of the sample study is how it can most effectively match the reporting cycle of the ANR. Since the sample study will have limited ability to function in an individual diagnostic manner, it is not necessary to leave enough time within a school year for corrective action. Testing towards the end of the school year should mean that students have had greater exposure to language learning, and have covered a greater proportion of the material included in the test. Experience suggests that schools find it inconvenient to participate in research projects in the last term of the school year, as there are many other competing demands on their time. The early finish for Year 10 students in some states also needs to be taken into account. These considerations suggest that a time close to the end of Term 3 is the optimum, causing the minimum of disruption and presenting the maximum chances for outcomes.

Consideration also needs to be given to whether this is a one-off exercise or whether it will be repeated in subsequent years. Improvement can only really be measured over time with repeated measures, preferably on the same sample. Since we would not expect there to be rapid changes in population proficiency on a year-to-year basis (however measured), unless there was an extremely significant intervention, annual testing is probably not required. Testing every 2, 3 or 4 years is possible and satisfactory, with a three-year cycle most likely to be acceptable, given competing priorities for national testing. Funding for the first sample study would be possible from within the national projects component of the current NALSAS budget.
7.4. Costs of a sample study

Precise costs for a sample study for Asian languages are not easily quantifiable at this stage because there are few precedents to work from, and decisions about what to test, how to test, and who to test will dramatically impact on costs. The development, administration, marking and analysis of the sample study assessments should be put to competitive tender to obtain the best value for money and possibility of innovative methods (such as web-delivered interactive testing).

The major determinants of costs are:

- **Size of the sample.** Any requirement to provide analyses at disaggregations less than state level greatly increases the size of the sample needed. If more than one year level is tested, the size of the sample will be increased proportionately.
- **Complexity of assessment development task** (and extent to which negotiations between stakeholders are required).
- **Complexity of assessment methodology.** Costs are least when the test can be administered and delivered by the classroom teacher to an intact class, and greatest when external invigilators or examiners are required. Costs are least for machine-scored paper and pencil tests and greatest for performance–based assessments such as oral examinations.
- **Number of languages to be tested.** There has been some suggestion that a generic languages test may be possible, but it is more likely that, particularly with older age groups, (because of the greater emphasis on written language in the curriculum in Stages 4-6), separate tests will be required for each NALSAS language. The differences between scripted (such as Mandarin, Korean and Japanese Kanji) and non-scripted languages such as Indonesian may be such that separate test instruments are required. (Also, if more than one year-group is tested, unless a common test is used, each version of the assessment instrument developed will add to the overall cost).

Sample studies in other areas completed in recent years have been conducted for around $200,000 (total contract costs excluding GST). This suggests that a sample study for Asian languages should be able to be delivered for well under $300,000 for each language. Since much of the cost is in development or “set-up”, costs in subsequent years should be lower.

**Conclusion**

This paper has considered some options for developing an outcomes reporting framework for NALSAS languages. While each option has advantages and disadvantages, given the lack of any comparable data that demonstrates student proficiency in existence in the states and territories, the most likely way forward appears to be to gather new data from a sample survey. Some issues in designing and implementing a sample study have been considered that will help NALSAS Taskforce decision making about what to test (test content), what languages should be tested, how students should be tested (test processes), and when to test.
Appendix 1.

Consultation in preparing the discussion paper

NSW
Ms Karen White,
Manager LOTE
NSW Department of Education and Training

Victoria
Ms Helen Spargo (for Dina Guest)
Manager LOTE, ESL and Multi-Cultural Branch
Education Victoria

Tasmania
Ms Sue Tolbert
Principal Education Officer, LOTE
Tasmania Department of Education

Queensland
Mr Roger White
Manager,
LOTE, SOSE and HPE
Education Queensland

Australian Capital Territory
Ms Pam Rosser
ACT Schools Authority

Western Australia
Ms Elizabeth Bragg
WA Department of Education

Northern Territory
Ms Kathy Silburn
Manager NALSAS Programs
NT Department of Education

South Australia
Ms Helen Touloo
SA Department of Education, Training and Employment

AFMLTA
Ms Angela Scarino
Director, Research Centre for Languages and Culture Education
University of South Australia

Commonwealth Department of Education, Training and Youth Affairs
Mr Noel Simpson
Ms Grette Toner