SECTION ONE

AIM AND OBJECTIVE

1.1 Aim and objective

The aim of the project is to identify cost effective and statistically reliable means of establishing equivalences among State and Territory student learning outcomes at the primary level. The equivalences will be used for ongoing reporting of comparable State and Territory student learning outcome information for the government school systems.

It was the view of the Schools Working Group that among possible indicators of school student learning outcomes, results of State/Territory-wide assessment would be the most useable. This is so because such results are the data currently available on learning outcomes in all States/Territories; they have an official status; they are being collected at specific intervals and are, thus, readily available on an ongoing basis.

The objective of Phase One of the project, therefore, is

> *to develop a detailed methodology to establish equivalences among existing State/Territory tests to enable ongoing reporting of nationally comparable learning outcome information.*

The following are the main features of the project:

- Phase One of the project (the current study) deals with the development of a detailed workable plan for aspects of literacy assessed by States or Territories
- Phase Two will involve the implementation of the plan for forming equivalences in aspects of literacy.

1.2 Issues addressed

This final report of Phase One has been prepared by the Macquarie Consortium, the research team commissioned by the Review of Commonwealth/State Service Provision, established by Council of Australian Government (COAG) to undertake the project *A plan for development of nationally comparable school student learning outcomes through establishment of equivalences between existing State and Territory tests*. The report addresses the following tasks:

- Realising the objective of Phase One of the project within the context of the project as a whole
- Identification of comparable nationwide learning outcomes
- Detailed specification of the theory and implementation of the methodology proposed for Phase Two and subsequent phases of the project concerning other learning areas
- Presenting how learning outcomes can be equated for each State/Territory at present, and over time, and choosing one of these as the optimal
- Providing a detailed account of the sampling procedure proposed for Phase Two, including the overall sampling plan and the State/Territory specific variations
- Establishing the testing instrument(s) to be used in Phase Two for equating the State and Territory assessment instruments
- Producing the implementation plan for, and the timeline of, the proposed plan for Phase Two
- Initial projection of the outcome of Phase Two, including the nature of the equivalences, and the nature of the reporting scale.

1.3 Guideline for Phase Two work plan

In developing the work plan for Phase Two, the Consortium has been operating under the following guidelines:

- The least disruption possible to the States/Territories should result from the project
- The testing regime specified for Phase Two has to be compatible with current arrangements in the States/Territories.
- As far as possible, the testing program proposed for Phase Two should make full use of existing assessment infrastructure in the States/Territories.
- The model developed in Phase One will be transferable across all learning areas in the school curricula.
- The data analysis and the result reporting in Phase Two will be focused on State/Territory-wide findings and implications with confidentiality guaranteed for school and student specific information.

1.4 The focus of the methodology

The main focus of the methodology for the establishment of equivalences is on matching the various scales of the State/Territory tests. The equivalences to be estimated relate to the scales in the State/Territory tests, rather than an immediate comparison of student performance. The latter can follow from the representative sampling of students within states, and the performances of these students, once the equivalences have been established. The distinction is important for a correct understanding of the methodology proposed, particularly regarding sampling and test materials to be used.

Two key issues in establishing equivalences became evident to the Consortium: (i) that the States/Territories defined literacy in noticeably different ways with implications for any intended matching of performances to the National Profiles, (ii) that the assessment of literacy within each State/Territory had inevitable limitations of sampling and practicality.

The implication of the first issue is that the equivalences take on the feature of scaling as well as equating. In scaling, the scores on two or more tests which have inherently different content are made equivalent in terms of relative difficulty (for example, scaling of performances across different subject areas in tertiary selection); in equating, the scores on two or more tests which have inherently the same content are made equivalent. Thus, the equivalences established across States/Territories will in part reflect the same content, and in part reflect equivalent achievement, but on different content. The distinctions will need to be appreciated in the conduct, especially in the reporting and interpretation, of Phase Two.

The implication of the second issue is that the range of skills and capacities assessed by the different tests in the States/Territories does not reflect the entirety of the skills and capacities that are being developed in various schools. Again, the limitations will need to be appreciated in the conduct, and again especially in the reporting and interpretation, of Phase Two.

At the end of Phase One, the sampling plan has been developed, the tests to be included for equivalencing have been selected, the statistical model to be adopted has been defined and a detailed work plan has been specified. In Phase Two, the actual sampling procedure must be implemented; the test papers must be printed; the statistical model must be implemented both in terms of its statistical configurations and computing program compilation before actual data analysis. In a complex project like the present one and with the very tight deadline foreseen in Phase Two, all data entry and analysis will have to be thoroughly rehearsed.

Details of the methodology are found in sections 3 to 8 of this report.