

An Even Start: Innovative resources to support teachers better monitor and better support students measured below benchmark



Jocelyn Cook

Australian Council for Educational Research

Jocelyn Cook is a Principal Research Fellow and Manager of the Perth Office at the Australian Council for Educational Research. With a background in teaching English and Literature, she has extensive experience in system assessment and reporting. As the Manager of the Monitoring Standards in Education (MSE) team in the Department of Education and Training (DET) in Western Australia, she was responsible for all system assessment programs, overseeing a number of highly innovative and successful random sample projects that included Technology and Enterprise, Social Outcomes of Schooling, the Arts, and Languages Other than English.

From 1998 until 2006 she oversaw the introduction and transformation of the Western Australian Literacy and Numeracy Assessment (WALNA) program. Under Jocelyn's management, Western Australia was at the forefront of innovation and development in population testing in Years 3, 5, 7 and 9, as well as research into the system level assessment of writing.

Since joining ACER in 2007, she has directed the development of *An Even Start Assessments*, a suite of computer-based assessments designed specifically for use with students experiencing difficulty in acquiring literacy or numeracy skills. She has also undertaken consultancies on behalf of the World Bank providing policy advice to the governments of India and Bangladesh; conducted and reported on a review of the introduction of on-screen marking for the Curriculum Council in Western Australia; and managed item development for the National Assessment Program in Literacy and Numeracy (NAPLAN).

Abstract

Literacy and numeracy are foundation skills for a successful education and a productive life. Improved literacy and numeracy outcomes encourage higher school retention rates, lead to improved employment opportunities, enhance engagement in society and support future economic prosperity. Conversely, a range of research indicates that poor basic skills are associated with a life trajectory of disadvantage. Enhancing teachers' capacity to recognise the specific needs of those students with the poorest skills and to provide remedial help is at the heart of breaking the cycle of disadvantage.

An Even Start is a Commonwealth government initiative aimed to address the learning needs of students who require additional assistance to acquire satisfactory literacy or numeracy skills. The resource developed for DEEWR by ACER focused on: accurate diagnosis of specific needs; provision of useful advice to teachers and tutors on remediation of specific difficulties; and more precise and accurate measures of progress.

This paper traces the conceptualisation of *An Even Start Assessments*, illustrating how the instruments and support materials draw together the requirements of good measurement, useful diagnostic information, and accessible and relevant teaching strategies.

Background

In the 2007–08 Budget, the Australian Government announced funding of \$457.4 million over four years to provide \$700 in personalised tuition (a minimum of 12 hours) to students in Years 3, 5, 7 and 9 who do not meet national benchmarks in literacy or numeracy. The tuition assistance was provided through *An Even Start*

– National Tuition Program. The Australian Government's desired outcome for the program was a measurable improvement in the literacy or numeracy levels of eligible students. *An Even Start* – National Tuition Program was a large and complex program, completed in a changing political environment, weathering a total of three federal Ministers as well as a change of government.

A key feature of the program in its initial conception was the provision of one-to-one private tuition managed by a state-based Tuition Coordinator. This was broadened in its second year of operation (2008) to include school-based tuition that could be delivered either: one-to-one; to groups of up to five students; or online. A requirement of *An Even Start* program was to give accurate diagnostic information and resources to tutors to support them in providing appropriate instruction for eligible students. The Australian Government also required from *An Even Start Assessments* the capacity to evaluate the success of the initiative.

A need for better targeting

It is widely acknowledged that state and territory tests, conducted until 2007, as well as the NAPLAN instruments, provide measures of ability with large measurement errors for students at the extreme ends of the scale. This is not a criticism of the quality of these tests; rather an observation on the measurement properties of any test designed to measure a wide range of abilities. This means that, for students at the lower extreme of the scale (typically with only one or two questions correctly answered), there is very little information on which to base the estimate of ability. Tests targeted for the general population consequently provide quite limited diagnostic information for very weak students.

The assessment instruments developed by ACER for *An Even Start* have therefore been carefully targeted to provide the finer-grained information about the strengths and weaknesses of those students reported below benchmark. The instruments have specific features designed to both support the student and facilitate accurate analysis of strengths and weaknesses.

Assessment features and design

Firstly, in all *An Even Start* assessment materials, particular care has been taken to ensure the assessment topics and activities are interesting and based on contexts that are not perceived as trivial or simplistic, even though the activities themselves are contained and directed. Secondly, the assessment tasks are designed to not only help the tutor gain insight into the particular needs of the student and provide multiple and independent observations on aspects of student achievement, but also to help the tutor establish a structured, purposeful and productive interaction with the student.

All numeracy instruments, for example, allow a level of teacher/tutor support (reading questions aloud where required) in their administration. The numeracy assessments also utilise a scripted 'interview' that allows the student to explain their mathematical thinking. These strategies are designed to limit the interference of reading skills with the diagnosis of numeracy difficulties. The numeracy items have also been constructed specifically to tap into common misconceptions that may be present in a student's mathematical thinking.

It is recognised that students who are struggling to develop adequate reading and writing skills are not well-served by conventional paper and pen tests. In some instances they may not have

established a strong awareness of sound/symbol correspondence and therefore are unable to effectively attempt a conventional reading or writing assessment. Older students who have experienced failure in reading and writing are often extremely reluctant to engage with assessment tasks and may exhibit passive or antagonistic behaviours. In the assessment of writing, students judged to be below the benchmark tend to produce very short texts, which provide extremely limited evidence on which to base decisions about attainment and intervention.

Innovations

An Even Start assessments of literacy directly address this issue by the inclusion of 'investigations': a series of short, highly-focused activities designed to give the tutor some specific insight into the particular difficulty a student may be experiencing. The program contains two sets of investigations. The first is the *Components of Reading Investigation* (CoRI), a series of small investigations to be conducted one-on-one with students in order to provide specific insight into the areas of difficulty experienced by those students who are not independent readers and are deemed to be below the Year 3 reading benchmark. The CoRI allows the teacher to focus on the student's phonemic awareness, understanding of phonics, vocabulary, and fluency. It is essentially diagnostic in purpose.

The second set of investigations is the *Investigations into the Components of Writing* (ICW) and these too are essentially diagnostic. They are specifically designed to give teachers more insight into the specific areas of difficulty for students struggling to develop writing skills. The areas for investigation in writing are sentence knowledge and control; punctuation; sequencing and cohesion; spelling/word knowledge; vocabulary; and ideas.

Software

An Even Start assessment instruments and support materials are provided on-screen through a purpose-built software package. The *An Even Start* assessment package contains materials targeted for use with students reported below the Years 3, 5, 7 and 9 benchmarks. The package has two key components: calibrated pre- and post-assessment tests for each year level that allow progress to be monitored; and links to resources or teaching strategies relevant to the particular point of need or weaknesses identified in the pre-assessment test. The post-assessments mirror the skills assessed in the pre-assessments although the post-assessments are a little harder overall so that progress can be measured.

Reading and numeracy pre- and post-assessments include both constructed response and multiple choice questions. Multiple choice question results are automatically uploaded when the assessment is done on-screen. Constructed response questions are scored using the marking guide, available from the system documents' Main Page. Once student scores have been entered into the software, a detailed diagnostic report on the student's performance is generated.

These reports show which questions the student answered correctly or incorrectly and which misconceptions may exist. Tutor resources, linked to each question or group of questions, are provided as website links in the software. These resources are as specific as possible. This means that if a student demonstrates difficulty, for example, with questions requiring control of place value, then the links are to resources that deal directly with supporting the development of that skill.

Measuring growth

The pre- and post-assessment items at each year level were calibrated to a common scale. These items were used to build a set of progress maps. These maps display skills typically associated with students working close to the benchmark level. The progress map contains sufficient detail to show the skills that need to be developed to make progress. To achieve pre- and post-test calibration, two sets of equating studies were conducted in 2008 in three states and across all sectors to establish year-specific and domain-specific scales. The scope and timeline of the original contract did not provide for equating to the national scale since the national scales were not constructed at the time *An Even Start* material was being developed. However, the instruments for *An Even Start* are designed to facilitate common-person equating, when national data are made available. This would allow for national benchmark locations to be applied to *An Even Start* scales.

Similarly, the scope and timeline of the original contract did not allow for the construction of two single scales for numeracy and for literacy. Again, the instruments were designed with items common to adjacent years, to facilitate development of single literacy and numeracy scales that will allow progress within and between year levels to be described.

Development potential

Although the tutorial system, for which this suite of assessment materials was initially designed will not continue, it is hoped that these materials will be made readily available for use with target students. As indicated, there is capacity to build on the *An Even Start* assessment tools that would enhance the usefulness of these materials. Should funding be available the following work is recommended:

1. equating the material to the national scales for reading and numeracy
2. calibration of the CoRI and ICW so that they become measures of the components of writing, rather than guides to early development
3. continued supplementation of the support material links, using jurisdictions' material when it becomes accessible.

Evaluation

The Department of Education, Employment and Workplace Relations (DEEWR) has commissioned an evaluation of *An Even Start* by the independent social research consultant, Urbis. The overall aim of the evaluation is to assess the success of the program in terms of its appropriateness, effectiveness and efficiency in achieving the objective of lifting the literacy (reading and writing) and numeracy performance of students who did not meet the national Year 3, 5 or 7 assessment benchmarks in 2007. As part of the evaluation, in March 2009, online surveys of a random sample of tutors and school coordinators involved in *An Even Start* have been conducted. The final evaluation report is due for submission in the middle of 2009.