

A case study of an Assessment Task Bank

STEPHEN H MOORE and MARIAN HARGREAVES – Macquarie University

ABSTRACT

An Assessment Task Bank (ATB) is an attractive option for a language-teaching institution with a criterion-referenced curriculum. In theory, with appropriate professional guidance, teachers could create and moderate tasks, which could then be stocked in the bank to be available for subsequent use by a wider pool of teachers in their classroom assessments. In practice, the process is not a simple one, especially if one of the goals is to have multiple tasks of good quality and approximately the same level of difficulty for each competency being measured. This paper outlines the processes involved in stocking an ATB with appropriate assessment tasks. It reports on the online ATB of the Australian Adult Migrant English Program (AMEP), highlights the major challenges that the task bank faces, presents some of the solutions that have been adopted, and addresses the complex issue of why the process for completing new tasks is inherently slow.

Please note the following abbreviations, which are used regularly:

AMEP – Adult Migrant English Program

ATB – Assessment Task Bank

CSWE – Certificates of Spoken and Written English

DIAC – Department of Immigration and Citizenship

LLNP – Language Literacy and Numeracy Program

LO – Learning Outcome

NWG – National Working Group

NSW AMES – NSW Adult Migrant English Service

Background

The AMEP provides English-language instruction to recently arrived immigrants in Australia through a competency-based national curriculum framework, which is known as the Certificates in Spoken and Written English (CSWE). This comprises four levels, from pre-beginner to intermediate levels: ‘The Certificates I–IV in Spoken and Written English provide [an accredited] framework for learners to develop English language and literacy skills and numeracy skills to undertake further education and training, seek and maintain employment, and participate in the community’ (NSW AMES 1992). The curriculum documents include an overview of the certificates, including the Preliminary Course; a brief outline of issues in assessment; and sections on institutional planning, language learning and syllabus design.

Each certificate level comprises modules that cover specific skills (for example, Certificate I, Module D: Beginner listening and reading skills for information texts). These modules include learning outcomes (for example, Certificate I, Module D, LO 1: Demonstrate understanding of a spoken information text). The modules may be core modules and therefore compulsory for the achievement of a certificate (for example, Certificate I, Module A: Beginner learning strategies). Indeed, all the modules for the Preliminary Course are core modules. Certificate I includes three core modules, while both Certificate II and Certificate III each have one core module. The CSWE 2008 now includes 128 learning outcomes, 33 of which are in core modules (NSW AMES 2008).

An online ATB was first developed in 2002 for the CSWE curriculum to standardise assessment procedures nationwide. This ATB is essentially a collection of tasks held in a secure location and to be used exclusively by AMEP teachers for assessing achievement in the CSWE curriculum. Its aim has been to include only tasks that have been trialled with learners to ensure that (1) the tasks assess what they are intended to assess, (2) they are of the right level of difficulty for AMEP learners, and (3) each task in a

given competency is of equivalent difficulty. Another function of the task bank is to provide models of tasks that teachers can use as blueprints for developing their own tasks, which can then be submitted to the ATB administrator for processing. These aims all respond to the difficulties in ensuring the comparability of teacher-developed assessments, as noted by Brindley (1994). Although the authors are aware of the existence of 'in-house' task (or test item) banks, their review of the literature could find no published accounts of any ATBs other than that of the AMEP. Thus, this paper hopes to contribute significantly to an under-researched and under-reported area of language assessment.

Introduction

This paper provides an update on previous research reported in relation to the AMEP's ATB. Brindley (2001a: 401) introduced the notion of 'a bank of fully-piloted exemplar tasks', while Slatyer (2003) described its further genesis and eventual implementation. The current paper discusses the ATB's subsequent development and some of the key issues arising in practice during the period to 2009.

The online ATB has been operating since 2002, and allows AMEP teachers to access and download an assessment task for a competency they have been teaching (known as a learning outcome in the CSWE curriculum). The tasks that are in the bank have undergone a rigorous vetting and development process, resulting in tasks that are quite superior to what untrained teachers could create on their own. This paper, therefore, seeks to shed light upon the processes and constraints in bringing suitable tasks into the bank, and to make some suggestions for the way forward. While some of the constraints are specific to the AMEP's ATB, the processes are applicable to any ATB project concerned with controlling the quality of assessment tasks and test items.

A brief history of the AMEP's ATB

Feez (2001) provides an enlightening account of curriculum evolution in the AMEP, reflecting on its first 50 years of existence. Against this historical backdrop, Slatyer (2003) presents a good overview of changes in assessment practices in the AMEP that occurred in 1993, when the curriculum framework became outcomes based. The rationale for setting up an ATB came from the findings of a series of research projects from 1996 to 1999, led by Geoff Brindley and Gillian Wigglesworth at the AMEP Research Centre. These projects tried to identify which features of tasks impacted on the performance of learners. A range of tasks was developed in which different features were manipulated (either task characteristics or task conditions), and the tasks were trialled in AMEP centres around Australia. The results of the research suggested that certain features of the tasks (such as the type of question or the speech rate in listening tasks) made them more or less difficult for learners (see Brindley 2000; Wigglesworth 2000; Brindley and Slatyer 2002; and, more recently, Brindley et al 2007).

It became clear from this research that there was a great deal of variability in the tasks that were being used for the assessment of the competencies in the CSWE curriculum, and that this would lead to unfair assessment. In other words, one learner might face a more difficult task than another learner of comparable ability, even though they had studied the same competencies. It also emerged (see Malcolm 2001, cited in Slatyer 2003) that teachers found designing assessment tasks to be difficult (ie time consuming and burdensome) and that they would much prefer to have access to a set of ready-made assessment tasks. Slatyer (2003) also notes that in order to satisfy these needs, new research projects, which took a two-pronged approach, were set up: (1) to conduct professional development workshops to train teachers and professional development personnel in assessment task design; and (2) to develop and trial a set of tasks for the creation of an ATB that could easily be accessed by teachers. The participants in the task design workshops would develop the assessment tasks for the task bank. Thus, a National Working Group (NWG) of AMEP staff and teachers was set up and trained in assessment task development.

The original goal was to have the NWG develop all tasks needed in the ATB, either in collaboration with other NWG members or by conducting professional development workshops with teachers at their own institutions. The NWG would work with tasks (submitted by AMEP teachers nationwide) by vetting them (ie selecting the most suitable ones for inclusion in the task bank) and modifying them (ie to comply with validity and reliability issues). Once the tasks had been modified, the NWG members piloted them with a small number of learners in their own centres.

The purpose of the pilot was to:

- detect any major faults in the tasks
- make sure the tasks produced performances that complied with the performance criteria specified in the CSWE curriculum documents
- check that the conditions of administration were consistent
- ensure that the tasks were of the right level of difficulty for the learners.

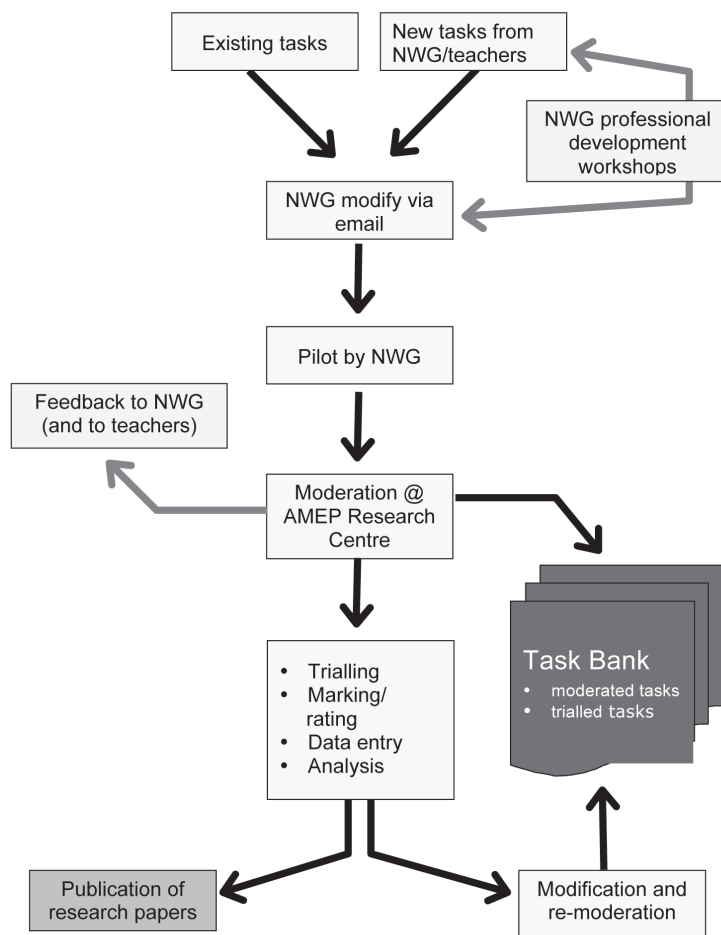
The NWG then modified the tasks based on feedback from the piloting process. After these modifications, the tasks were trialled with large numbers of learners around Australia. The trialling had similar objectives to the pilot study but with larger numbers to enable the running of statistical tests on the results. The purpose of these trials was to ensure that each learning outcome had at least one task with known measurement properties (through using Rasch analysis), which could therefore act as an ‘anchor’ for that particular learning outcome Brindley (2001a: 401). At this stage, some of the tasks were rejected or modified once again and re-trialled. Once the results of the trialling had been analysed, the tasks were calibrated (ie they were ordered in terms of their difficulty). Only at this stage were they finally ready to be put in the task bank.

From this description of the original process of developing assessment tasks for the task bank, one can easily see just how labour-intensive and time-consuming the process was. In 2004, however, following funding cutbacks to the NWG, the task development process itself had to be modified to make it more streamlined and less dependent on the NWG’s time. The revised process is set out in the following section.

The current procedure for tasks to enter the ATB

Figure 1 illustrates the process by which tasks currently enter the system and, ultimately, may end up in the ATB.

Figure 1: Overview of task development process from beginning to end



Tasks are submitted electronically to the ATB administrator by members of the NWG or by AMEP teachers independently. The ATB administrator formats the tasks into the house ATB format, checks that there are no insurmountable copyright issues, and ensures that all elements of the task are present (for example, that an answer key has been provided). Instructions are rephrased into the standard for the task, and questions checked to remove duplicates, dichotomous questions and any other questions that may be inappropriate for the criteria. The new task is then logged and forwarded to the NWG via a NWG email list. The NWG members work collaboratively (by email) and check the tasks for suitability for the particular learning outcome. The tasks may or may not be piloted at NWG member institutions. The NWG may suggest modifications to the task or recommend (with reasons) that it be discarded as unworkable. The tasks returned to the ATB administrator by the NWG are then further moderated by an 'expert' panel consisting of language assessment specialists at Macquarie University, AMEP Research Centre staff and AMEP teaching staff. After this moderation session, the tasks are deemed fit to enter the ATB, labelled as 'moderated' tasks, and are available for all AMEP teachers to use in assessing their students. These tasks are then scheduled for trialling nationwide. Once the results of the subsequent trialling have been analysed, the tasks are calibrated by the AMEP Research Centre and then are finally ready to be put in the task bank as 'trialled' tasks. To sum up, the current procedure for navigating tasks into the ATB is as follows:

- tasks are sourced from teachers, NWG members or AMEP Research Centre staff (these tasks have, preferably, already been piloted)
- tasks are critiqued and modified by members of the NWG through professional development workshops or through a group email list
- final versions of tasks are prepared by the ATB administrator
- tasks are moderated by an expert panel consisting of language-testing academics, AMEP teaching staff and AMEP Research Centre staff
- moderated tasks are made available in the Task Bank
- moderated tasks go to national trial
- trialled tasks are modified if necessary
- modified trialled tasks are made available in the Task Bank.

The current process of developing assessment tasks for the ATB, though less rigorous than the original procedure, is still heavily labour intensive. Any further shortcut of the process could easily result in more tasks entering the ATB but they would be of relatively poorer quality. It must be acknowledged that language assessment is inherently a very time-consuming enterprise (Alderson, Clapham and Wall 1995; Bachman and Palmer 1996), and it is not uncommon for resources to be unwittingly spent developing well-intentioned but seriously flawed assessment instruments (Hughes 2003).

In addition to the time-consuming processes just described, since the establishment of the task bank, the CSWE has been revised twice – first in 2002 and again in 2007, with accreditation in January 2008. These revisions entailed the ATB mapping the new curriculum learning outcomes onto the existing ones, which therefore required the time-consuming re-evaluation of every online task.

Current completion status of the AMEP's ATB

The ATB deals with the first three levels of the CSWE and aims to provide at least three tasks for each learning outcome that requires formal tasks for assessment. It is worth noting that in the CSWE curriculum, some learning outcomes require 'guidelines' rather than formal tasks. These guidelines allow more flexibility for creating tasks suitable to the particular context of a given class (for example, Certificate II, Module A, Learning Outcome 1: Develop a learning/training plan). The curriculum also includes a number of modules for the assessment of numeracy (for example, Certificate II, Module N: Post-beginner numeracy skills for dealing with basic calculation and measurements in familiar contexts). In order to meet government requirements and prioritise resources, the task bank has a current focus on literacy.

The current completion status of the ATB, for formal literacy tasks for the CSWE 2008, is set out in Table 1.

Table 1: Status completion of the AMEP ATB, 10 October 2008

Tasks	Pre-CSWE	Cert. I	Cert. II	Cert. III	Total
Total number of LOs in curriculum	21	32	34	41	128
No. LOs requiring formal tasks	3	25	30	31	89
No. LOs suiting alternative assessment	18	7	4	10	39
No. numeracy LOs	3	9	8	5	25
No. prioritised LOs *	3	20	22	26	71
LOs with 1+ task online	2	9	6	7	24
No. LOs with 2+ tasks online	0	6	4	6	16
Coverage of LOs with 1+ task in ATB	10%	28%	18%	17%	18% average
Coverage of prioritised LOs with 1+ task in ATB	–	–	–	–	76.06%
No. tasks under development (not online)	17	42	68	48	175
No. tasks online for assessment	2	64	34	29	129
Total number of tasks currently in the ATB					304

* prioritised learning outcomes (LOs) = LOs that are not numeric and require formal assessment

Since the accreditation of the current CSWE curriculum, in January 2008, the NWG has evaluated all of the tasks that were online for the CSWE 2003 curriculum and has updated many of those that only required minor modification; there is now a total of 129 tasks online for assessment. (This compares with the 221 that were available online at the end of 2007 for the CSWE 2003.) Despite the increase in the number of learning outcomes, from 90 in the CSWE 2003 to 128 in the CSWE 2008, 76.06% of the currently prioritised learning outcomes have at least one task online. In 2008/2009, 212 new tasks were sourced, and the task bank currently (April 2009) has a total of 304 tasks on its books.

Creating appropriate criterion-referenced assessment tasks

Having reviewed the broad processes by which assessment tasks flow into the ATB, and noted the current completion status of the ATB, it is now worth focusing in more detail on the issue of test development and how task writers meet curriculum specifications against which individual assessment tasks are designed and need to be judged. This, too, is a time-consuming activity.

The fundamental testing principles of validity, reliability and practicality (Alderson, Clapham and Wall 1995; Hughes 2003; Brown 2005) are no easier to put into practice in an ATB than in any other testing scenario. Validity, which refers to testing that which you think you are testing and nothing else, can only be achieved if the assessment task closely relates to the specifications set out in the learning objective documentation. Validity is enhanced through trialling tasks (and making appropriate adjustments based on the results of trialling) and rating them for difficulty, hence the importance of this aspect of the task preparation procedure. The CSWE curriculum framework specifies assessment task characteristics. Therefore, an assessment task developer must look closely at a particular learning outcome and design a task that fits the requirements of the specifications. This in itself is not an easy undertaking, although it might appear to be relatively straightforward. For example, a reading text of 120 words is rather short if it is to have, say, six question items drawn from it. In such circumstances, it is very easy to have questions that overlap one another, thereby confusing precisely what is being tested by each item. Similarly, multiple choice items can be very difficult to design when a fourth distractor is needed but is simply not forthcoming from the limited available text or context. In addition, a task that aims to test one particular skill, such as listening, could place heavy demands on another skill, such as writing. Listening tasks should, therefore, have the option of an oral response and/or require a minimal written response.

Reliability, which refers to consistency in measuring the learners' performance of a task, likewise is enhanced when tasks are moderated, trialled and adjusted for anomalies. Reliability is also improved

through provision of clear and consistent instructions for test administration. Having created an appropriate assessment task for the ATB, it needs to be properly formatted and ‘packaged’. The core components of the task bank are, of course, the assessment tasks. However, these tasks are issued in a ‘set’ that also includes (1) a procedure for administering the task (the CSWE framework also specifies assessment performance conditions), (2) a task sheet to be given to the learners, and (3) an answer key for listening and reading tasks. An example is given in Appendix A. The reason for including these materials with the task sheets is that an essential part of the task is the way it is administered. By including instructions on how to administer the tasks, the variability in the assessment can be reduced. The reason for including answer keys with the receptive skills tasks is also to ensure that there is as little variability as possible in the way the tasks are marked. Research has shown that teachers often use different marking strategies, with some teachers being more severe or more lenient than others (see, for example, Smith 2000; Wigglesworth 2000; Brindley 2001b). The answer keys ensure that every answer is marked using the same range of responses.

Practicality, which focuses on the best use of limited resources of time, money and expertise, is improved when needs and output are prioritised. For example, the AMEP’s ATB has focused first on compulsory learning outcomes, rather than elective learning outcomes. Much of the criticism over the lack of tasks in the ATB is in regard to elective learning outcomes that have been allocated a low priority and, therefore, have not had tasks developed for them. The current ATB has focused on literacy tasks, for lower levels of competency and work-related topics, which has inevitably resulted in a lack of tasks for other learning outcomes. It should be noted, however, that all tasks, where available, are presented in a format such that they are ready for immediate use and require minimal preparation on the part of the teacher.

Trialling

Trialling is perhaps the most time-consuming yet important phase of the task development process, and is worth considering in detail here. If tasks are not trialled by actual learners, then a task bank could potentially be full of flawed assessment tasks. It is only through properly conceived and rigorous trialling of tasks by AMEP learners that weaknesses in tasks can be identified and remedied. Indeed, were it not for the trialling stage, the ATB would be relatively easy to complete quickly. Flaws that have only emerged from trialling, in seemingly excellent tasks, have included topics that were not relevant to the students; ambiguity in both the questions and the answers; ageing information (such as the price of buying a car); and confusing information (such as hard-to-pronounce street names).

As shown in Figure 1, once assessment tasks have been moderated by the panel of experts, they are ready to be trialled. Trialling is prepared and coordinated by the AMEP Research Centre staff with the help of the NWG members, who coordinate the trialling locally. The phases of trialling are as follows:

- preparation of trialling materials (tasks, procedures, audio-cassettes for speaking learning outcomes, coding system)
- sampling (random allocation of tasks across centres and states)
- administration of the trialling materials (including returning results to the AMEP Research Centre)
- marking/rating of trialled tasks
- subsequent quantitative analysis of trialled data.

The success of the trialling process depends on several factors, not least the willingness of teachers to trial these tasks. Trialling is time consuming for teachers and no compensation is usually given for the extra time spent by those who agree to participate. Two national trials were held for the 2003 version curriculum tasks, the first being a small trial of three learning outcomes in late 2003. The second was a far larger trial covering 13 learning outcomes and took place in early 2004. More trialling was done in 2006 and focused largely on topping up the responses for learning outcomes trialled in 2004, and on the new tasks for assessing citizenship at CSWE Level I. Trialling showed that the citizenship tasks were seriously flawed. They were subsequently revised and successfully re-trialled in 2007. In total, tasks from 24 learning outcomes in the ATB were trialled in 2007.

Before any statistical analyses can be performed, the research design for the trials requires 100 samples for each of the ‘control’ tasks, and 20 samples per ‘experimental’ task. This is the minimum number of samples for the statistical analyses to provide significant results. Problems in achieving acceptable sample sizes can arise (for example, when fewer students attend classes for assessment than expected), and this can impact adversely on the supply of trialled tasks available for placing in the ATB. Nevertheless, an analysis of the trialling of reading and writing tasks has been completed (see Bowering and Lloyd 2006; Lising 2007). The analysis confirmed several issues of earlier research (such as factors affecting difficulty and the importance of consistency), highlighted the important role that the conditions of assessment play in a learner’s achievement, and reconfirmed the quality of the tasks in the ATB.

Key challenges to the development of the ATB

In addition to the significant challenges already highlighted, which any substantive task bank might face (ie matching the specifications of the criteria, the rigorous process of task development and the completion of national trialling), the ATB is challenged by a number of other factors.

Revision of the CSWE curriculum

Since its inception, the CSWE curriculum has been revised twice – in 2003 and 2008. On both occasions this involved significant and substantial changes to the content of the curriculum. In turn, this meant profound changes to assessment requirements, including an increase in the number of learning outcomes overall. Revision of the ATB to comply with the requirements of the new curriculum therefore involved:

- making large-scale changes to all existing tasks to conform to the new specifications
- developing a large number of new tasks for new learning outcomes
- adapting all procedures for the administration of revised tasks
- developing new procedures for new learning outcomes
- moderating all tasks (new and revised)
- large-scale trialling of all tasks (national trialling involving sampling, administration, marking/rating and analysis).

In addition to the tasks associated with the revision to the CSWE curriculum, an increase was sought in the number of assessment tasks for each learning outcome in order to give teachers some choice in selecting appropriate assessments for the topics and content of their classes and the particular profile of the learners in their classes. Accommodating this change meant that while there might be more assessment tasks available for a particular learning outcome, they would be generated at the expense of broader coverage of assessment tasks across other learning outcomes.

However, the revision of the curriculum created an enormous impetus among the AMEP’s service providers to develop new tasks. This included not only tasks for the new learning outcomes but also checklists for observational assessment. This required service providers to revisit the criteria of modules and learning outcomes in order to become familiar with changes to old requirements and accommodate the new developments. The overall approach of the 2008 curriculum is also more holistic than previously and the ATB is in an ideal position to provide models and guidelines for teachers seeking to integrate various learning outcomes.

Change of focus

A further issue impacted on ATB efficiency when, in late 2004, the Department of Immigration and Citizenship confirmed that the ATB should focus on the lower levels of the CSWE curriculum because the learner profile of AMEP clients was increasingly at lower levels of English proficiency. This shift in policy meant that fewer assessment tasks at higher levels could be accommodated by the NWG and ATB administration. Consequently, some work already underway with higher-level tasks had to be set aside as the ATB team adjusted its focus to lower-level assessment tasks.

Funding

Although the ATB's annual funding has been secure for a number of years, its budget only allows for an ATB administrator to work three days per week. Likewise, following budget cutbacks in 2004, the amount of time that NWG members are funded to work on NWG matters is very limited, averaging only a few hours per week. Thus, the heart of the ATB operation is run by part-time staff, and this seriously impacts on the capacity and, to some extent, efficiency of the enterprise.

However, these staff, through the task bank's NWG, are directly connected to the learners in the AMEP and to the teachers who use the ATB for their assessment. The ATB is therefore highly attuned to the needs of these learners and teachers, and can respond directly to quite specific requests. New tasks in the early stages of development can be piloted with minimal delay with students at the appropriate level of learning, and the feedback channelled immediately into the improvement of the tasks. As part of the AMEP Research Centre, the ATB coordinator is also able, through AMEP national forums and AMEP conferences, to meet personally with members of the NWG and re-establish connection with the members of this national network. This facilitates the delivery of updates on the status of the task bank and is particularly effective in getting feedback directly from teachers in the field. Such national gatherings also offer an opportunity for the ATB coordinator to meet other members of the AMEP (ie teachers and administrators) who may not be familiar with the task bank or the resources that it provides.

The limited annual budget of the ATB is also a major challenge to the moderation stage of task development (for which costs are not insignificant), as it is vital that testing-specialist academics and AMEP teaching staff attend. This has resulted in the moderation process becoming a 'bottleneck' in the overall process of preparing tasks for the ATB, as members of the moderation panel can only meet five times a year. Solutions have included splitting the panel into two comparable groups for the moderation of tasks. This has had an undesirable consequence, however, as each member of the panel invariably has his or her own useful insights to contribute for every task under development. A more successful strategy has been to distribute to the whole NWG the list of tasks selected for moderation at least two weeks before the meeting. This facilitates the provision of directly relevant feedback for those tasks and reduces the time needed for discussion of each task on the day of moderation. Nevertheless, there is still a large backlog of tasks under development and the progress of completing the task bank continues to be slow, typically averaging 12 tasks per moderating session.

Marking and rating of many hundreds of trialled tasks is also very time consuming and costly: marking has to be moderated and performed manually at the AMEP Research Centre; rating involves training and standardising raters, who then must manually rate hundreds of spoken recordings or written scripts. A pool of trained raters has now been established, and although it is difficult to organise a session, a minimal amount of retraining is now required to ensure inter-rater reliability.

Technology

Technology is both a challenge and an opportunity. Many teachers have limited access to the Internet, to colour printing, or to digital recording technology. However, as the 2008 revision of the CSWE curriculum also coincided with a re-development of the website of the AMEP Research Centre, of which the ATB is a part, it was possible to present teachers in the AMEP with a new website. This enabled the presentation of tasks in a distinctive format that set them apart from the 2003 curriculum, which many service providers continue to use until its discontinuation at the end of 2008. In effect, the task bank was 'spring cleaned' and was able to put into effect many of the lessons learned through the trialling of tasks developed for the 2003 curriculum. For example, following the CSWE revision for Certificate I, Module F, LO 2: Give a short spoken description, images were reassessed and, based on the feedback from trialling, several images were culled from the task set and several new colour images were sourced.

The website is also a vehicle for new technology. While recordings of texts for listening tasks have been, and continue to be, made available on CD, digital recordings can also be downloaded directly from the online task bank. These recordings, which are of high quality using a variety of clear, fluent speakers, make a significant contribution to maintaining the high standard and consistency of the listening tasks.

Professional development

A further result of the new AMEP RC website was the revision of the old Professional Development Kit for assessment tasks. This was a kit developed for teachers to use for professional development workshops on the development of assessment tasks. It was very popular and was used quite extensively. However, with the new curriculum, revision was necessary and it provided the opportunity to turn the kit into a self-access resource for teachers. This resource relates directly to the 2008 curriculum and provides revised examples and activities, together with checklists for developing, piloting and evaluating tasks.

Independence

As an enterprise that is independent of the NSW Adult Migrant English Service (NSW AMES) – the author and publisher of the CSWE curriculum – the ATB is in a position to both support the NSW AMES interpretation of the criteria for the learning outcomes (as provided through sample tasks for the CSWE) and to augment it. The ATB aims to provide tasks that are both a set of materials for use in assessing learners and models for teachers to follow in developing their own tasks. While the ATB collaborates where possible with the NSW AMES, the ATB tasks, nevertheless, give teachers a greater choice of available tasks to choose from, and often a different approach to fulfilling the requirements of the specification. The 2008 CSWE encourages alternative and integrative assessment; the ATB aims to support this autonomy of teachers in the AMEP.

Workshops

An additional response to the challenges faced by the task bank has been the re-establishment, in 2007, of an annual three-day workshop for the NWG. This brings together all the members of the NWG for a full three days, during which issues affecting the task bank can be directly discussed and resolved. With the decision, in 2007, to discontinue the annual AMEP conference, which previously provided the opportunity for the members of the NWG to meet face-to-face and to run a short, half-day workshop, this three-day workshop has been enormously valuable and successful. All the NWG members (or their representatives) were able to attend the workshop held in July 2007 and were able to discuss issues that would enhance their knowledge of task writing, improve the process of task development, clarify their roles as members of the NWG, discuss issues such as the accreditation of a new curriculum for the CSWE, and write and moderate new tasks. In January 2008 the NWG was able to address the issues of the newly accredited CSWE curriculum, map the old curriculum to the new CSWE, and provide AMEP teachers with tasks for assessment that met the new criteria of the curriculum.

Conclusion

The AMEP's ATB is a unique resource in terms of setting a national standard for adult English as a Second Language assessment in a competency-based curriculum. Though simple in principle, creating and stocking an ATB is a major undertaking involving the coordination of complex processes and very considerable amounts of time and money. This paper has sought to clarify what these processes are, how they are constrained by limited resources, and how they are dependent on the active participation of AMEP teachers and the ATB's NWG. Developing and maintaining the ATB remains an exciting and worthwhile endeavour. The challenges that it faces are significant, but the responses from the teachers for whom it is provided give unequivocal evidence of its success, despite its perennially incomplete status. The 2008 revision of the curriculum has ensured that the task bank remains in the forefront of educational pedagogy and reflects current research in assessment of a criterion-referenced curriculum. Tasks continue to be developed that are relevant to both the learning needs of adult immigrants and to the testing needs of AMEP teachers, who require valid and reliable assessment instruments.

Acknowledgment

The authors would like to acknowledge the work done by Geoff Brindley and Helen Slatyer in setting up the online ATB and in unpublished documents from which portions of this paper have been drawn.

References

- Alderson, J. C., & Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bowering, M., & Lloyd, R. (2006, October). *The jury is out: Reading trial analysis*. Presentation at AMEP National Conference, Perth, Australia.
- Brindley, G. (1994). Competency-based assessment in second language programs: Some issues and questions. *Prospect*, 9(2), 41–55.
- Brindley, G. (2000). Task difficulty and task generalisability in competency-based writing assessment. In G. Brindley (Ed.), *Studies in immigrant English language assessment*. Vol. 1 (pp. 125 – 157). Sydney: NCELTR.
- Brindley, G. (2001a). Outcomes-based assessment in practice: Some examples and emerging insights. *Language Testing*, 18(4), 393–407.
- Brindley, G. (2001b). Investigating rater consistency in competency-based language assessment. In G. Brindley & C. Burrows (Eds.), *Studies in immigrant English language assessment*, Vol. 2 (pp. 59 – 80). Sydney: NCELTR.
- Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 19(4), 369–394.
- Brindley, G., & Hargreaves, M., & Moore, S., & Slatyer, H. (2007). *Exploring reading assessment and item difficulty*: (Unpublished research report) AMEP Research Centre, Macquarie University.
- Brown, J. D. (2005). *Testing in language programs*. New York: McGraw-Hill.
- Feez, S. (2001). Curriculum evolution in the Australian Adult Migrant English Program. In D. R. Hall & A. Hewings (Eds.), *Innovation in English language teaching* (pp. 208–228). London: Routledge.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Lising, J. L. (2007, May). *The jury is out (again): Findings from the Assessment Task Bank writing trial*. Presentation at AMEP National Forum, 'Learners in transition', Adelaide, Australia.
- NSW AMES. (1992). *Certificates I–IV in Spoken and Written English*. Sydney: NSW Adult Migrant English Service.
- NSW AMES. (2008). *Certificates I–IV in Spoken and Written English*. Sydney: NSW Adult Migrant English Service.
- Pierce, L.V. (2007, March). *Classroom-based assessment of oral language*. Presentation at 41st TESOL Convention, Seattle, Washington.
- Slatyer, H. (2003). Responding to change in immigrant English language assessment. *Prospect*, 18(1), 42–52.
- Smith, D. (2000). Rater judgements in the direct assessment of competency-based second language writing ability. In G. Brindley (Ed.), *Studies in immigrant English language assessment*. vol. 1 (pp. 159 – 189). Sydney: NCELTR.
- Wigglesworth, G. (2000). Issues in the development of oral tasks for competency-based assessments of second language performance. In G. Brindley (Ed.), *Studies in immigrant English language assessment*, Vol. 1 (pp. 81 – 123). Sydney: NCELTR.

Task Set for 'Clifford Street Medical Centre'
Certificate II Module D Learning Outcome 1
Demonstrate understanding of a spoken information text



This task set should contain the following:

- Teacher instructions
- Tapescript
- Learner response sheet
- Answer key

*Digital recording now
available to download*

Teacher Instructions

Administering the task

- Cue the recording (if available) to the beginning of the task.
- Explain to learners what they will have to do, ie complete the sentences or answer the questions with information from the text.
- Allow learners to use dictionaries.
- Give the learners the response sheet.
- Inform learners of how many items must be correct in order to achieve the Learning Outcome, including any compulsory questions.
 - give learners time to read the questions
 - start the recording
 - first hearing of the listening text
 - silence for checking
 - second hearing
 - silence for checking
 - third hearing
 - silence for final checking
- Collect all response sheets from the learners.

Marking and feedback to learners

- When marking the learners' responses please adhere strictly to the answer key provided in the task set. If, however, you find a legitimate response that has obviously been overlooked, please mark it correct and let us know.
- As it is critical to keep all documents in the task set secure for future assessment purposes, please DO NOT hand out the answer key after marking or allow learners to keep their corrected response sheets when giving feedback.

Transcript

Clifford Street Medical Centre

“You have reached Clifford Street Medical Centre.

The Medical Centre is now closed.

Our opening hours on weekdays are 8am – 6pm. On Saturdays we are open from 10am – 4pm. We are closed on Sundays.

In an emergency you can contact the doctor on call on 0407 003 111, That’s 0407 003 111.

Or you can go to the Emergency Department of Victoria Hospital which is located on Park Road, Liverpool, across the road from Woolworths.

As of next month patients of the Medical Centre will be seen on a first come, first served basis. This change is due to the difficulty in keeping appointments. Patients will now be seen in the order that they arrive at the Medical Centre. Our busiest times are before 9am and after 3pm on weekdays. To avoid a long wait, we advise that wherever possible, patients should come during school hours or phone ahead to check on waiting times.

We would like to remind patients that we no longer bulk bill. Bills must be settled before leaving the Medical Centre by cheque, credit card or cash.”

CSWE 2008 Certificate II**Module D Learning Outcome 1: Demonstrate understanding of a spoken information text****Clifford Street Medical Centre**

Name of learner _____ **Date** _____

Answer the questions while you listen. You will hear the recording three times.

1. The opening hours for the Medical Centre on Saturdays are from ____ to ____.
2. The phone number of the doctor on call is _____.
3. The Emergency Department is opposite _____.
4. The busiest times at the Medical Centre are before _____ and after _____ during weekdays.
5. Patients are advised to arrive during _____.
6. You can pay your bill by cheque, _____ or _____.
7. Tick (✓) the correct answer.
This text is a:

recorded telephone message

radio advertisement

public announcement

telephone conversation

For teachers only

Outcome of task _____ Date _____

Name and signature of assessor _____

Answer key

In order to achieve the learning outcome, learners must identify the context/topic (Q7) and identify explicitly-stated information.

Note that sentences need not be grammatically correct or in sentence form but errors must not interfere with meaning.

1. 10 (am) to 4 (pm)
2. 0407 003 111
3. Woolworths
4. 9am /3pm
5. school hours
6. credit card, cash (*both must be correct for one mark*)
7. recorded telephone message