btl learning & assessment

AlphaPlus

Speed and Accuracy in Skills Assessment

White Paper

October 2009

Authors:
John Winkley - Alpha*Plus* Consultancy Ltd
Ian Brumfitt - BTL Group Ltd

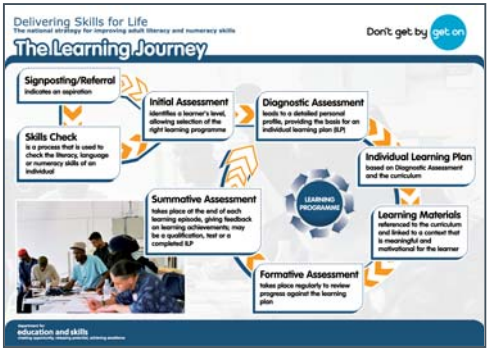
Introduction

This White Paper discusses the techniques and benefits of computerised skills assessments, which BTL and Alpha*Plus* have pioneered over recent years within UK post compulsory generic skills training such as Key Skills, Skills for Life and Functional Skills. In particular, it looks at the enhanced formative assessment opportunities deriving from computerised approaches and, as a result of technical and assessment design innovations, how these benefits can be realised in bespoke assessments customised for particular learning settings.

Learner and tutor expectations of diagnostic and formative assessment

Placement assessments are commonly used at the start of a learning programme to identify the best course for learners, and are as a result the learner's first experience of the learning programme. Although strictly these are more diagnostic than formative in purpose, for the learner at least, they must be quick, accurate, purposeful and engaging – delivering usable, reliable results in as short a time as possible and providing a learning experience suitable for the start of a learning programme.

The strategy in England for improving adult literacy and numeracy skills uses a pedagogy called "The Learning Journey", summarised in the diagram here. It involves learners moving from a preliminary skills need check (15 minutes, to establish a need) through initial assessment (25 minutes, to identify approximate learner level for course placement), and on to diagnostic assessment (1 hour+, to identify particular strengths and weaknesses at the individual skill/topic level in support of the Learning Plan – the first truly formative assessment in the programme.)



The Learning Journey Source: Department for Education and Skills

At the heart of all formative assessment is the need for the results of the assessment to be timely and actionable – formative assessment only works if learners and teachers are able to change their learning activity as a result of the assessment. For this reason, and recognising the time pressures on most learning today, developments since the introduction of "The Learning Journey" have focused on providing detailed, actionable feedback as part of all assessments, wherever they are used in the learning process. For example, even though the main purpose of initial assessment is to provide an assessment of level, these tools should also offer topic level feedback, and (as for all formative assessments) the opportunity for the learner to review their responses topic-by-topic, item-by-item, comparing them with the worked solution.

Sample literacy assessment feedback Source: *Skills For Health*

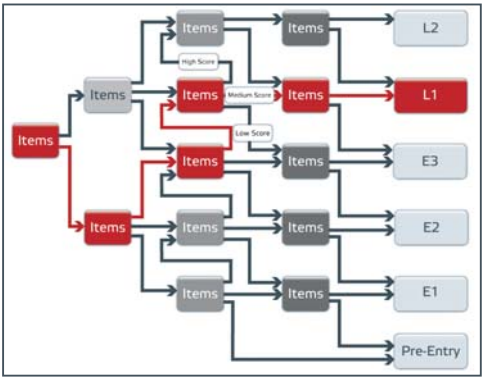
For learners returning to learning, the mediation of this assessment activity by the non-judgemental computer, and the availability of immediate feedback, are both powerful motivators – breaking down problems into manageable tasks, and also identifying strengths to build confidence. Assessments conducted in this way avoid the criticisms of "externalised over-assessment" as they become an integral part of the learning process – of personal value to the learner.

Speed and accuracy in skills assessment

Adult generic skills curricula cover a broad range of topics at a very wide range of levels. For example, the Skills for Life Numeracy Core Curriculum ranges from "counting to 10" at Entry Level 1 through to performing multistage numerical calculations at Level 2, and includes around 150 topics across the five levels. Of course, few assessments ever manage to test every topic in a curriculum (assessments are selective, due to time constraints and practical difficulties in assessing some skills, for example). But the challenge to produce a skills assessment which is valid in terms of coverage, but only lasts half an hour or so, is substantial and means that a single assessment providing a reliable result across a range of levels is essential. For qualifications such as Functional Skills, there is an additional requirement for the assessment to be contextualised – meaning that the learner must deal with source material, inevitably increasing reading time.

GCSE-style examinations use a compensatory model for assessment – marks are achieved on items with varying difficulties, and then various "cut scores" are used to place candidates into one of a range of grades according to their score. While it might in principle be possible to place candidates into "levels" according to this process, the generic skills approach is that candidates have to demonstrate a broad range of competencies at a particular level – being above that level in some topics brings no additional benefit as far as competency at the level is concerned. Hence a compensatory approach to competency assessment presents significant challenges to measurement accuracy.

In practice, the only way to achieve a reasonable coverage in a single test for all levels, so as to be confident of making a level judgement, is to use adaptive assessment. Adaptive assessments select harder or easier questions based on the computer marking of the learner's responses to previous questions or clusters of questions. This technique can be used to ensure that most of the questions the learner attempts are roughly at their level – avoiding "wasted assessment time" asking questions that are far too hard or too easy. This approach reduces assessment time, and helps with the production of a "spiky profile" – a common type of learner skills profile where the learner is



Example of a typical algorithm for an Initial Assessment tool

significantly stronger in some skill areas than others. The other critical advantage of an adaptive test is that a single test produces a result across all levels – requiring no prior judgement about "which test/level to put the learner in for".

Of course, adaptive, computer-marked assessments also have the advantage of providing immediate results – reducing the marking workload, but also providing feedback to the learner straight after the test, before the questions are forgotten, thereby increasing the opportunity to make the assessment formative.

Our approach to adaptive assessment

A simplified version of a typical algorithm for an Initial Assessment tool is shown above.

The "snakes and ladders" structure is used to route learners upwards, downwards or straight on according to their performance at the end of each cluster of questions. In the "First Filtering", questions begin at Entry Level 1 and end at Level 2. Those candidates who score above the threshold mark for this section are routed upwards to attempt Level 1 and 2 questions. Those who do not are routed to Entry Level questions in the "Second Filtering", and so on. In this way it is possible for learners to climb to the highest level, even after a poor start.

The adaptive nature of the tools makes them less threatening than conventional tests for learners as they are always asked questions broadly at their own skill level. All of the questions are aligned to assessment criteria from the adult core literacy and numeracy curricula.

Assessing the curriculum effectively

Traditional assessments of numeracy and literacy have been limited by what can be tested easily on paper. Presenting questions on-screen provides the opportunity to increase the range of skills that are tested, and to test them in a more authentic way.

For example, listening forms an important part of the literacy curriculum. The simple use of audio in on-screen assessments enables powerful assessment of listening skills. Similarly, while it isn't yet practical for the computer to mark candidates' written prose, it is acceptable to assess spelling, punctuation and grammar skills, and other constructed response skills without resorting to multiple choice questions. All types of literacy assessment are improved by the use of realistic documents –

Sample literacy assessment Source: Skills For Health

with colour, pictures, video and sound as well as the written word. In mathematics, candidates are expected to produce graphs and charts, show their working and calculation steps, and use tools such as calculators and measuring devices in order to undertake tasks. In this way, by requiring candidates to be more 'productive' – generating answers themselves rather than picking from a list, the assessments provide a more authentic test of skills and process.

Our assessments include audio for almost all questions – reading out the question for the learner. This simple feature reduces a common problem experienced in generic skills assessment, where the candidate struggles to answer the question not because they lack the skills being assessed but because their reading skills aren't sufficient to understand the question.

It is also possible to provide learners with "hints" (with the penalty of a reduced tariff for the question) so that they may obtain an answer to a question on which they would otherwise have been unable to complete.

A high quality learner experience

Learners want assessments that are clear, quick and easy to use, innovative in design and stimulating to use. Adaptive tests present items at just the right levels of challenge, while the use of a range of innovative item types makes the assessment feel authentic and avoids the monotony of MCQs. Extensive trialling has shown that learners really value these approaches – this makes them more likely to keep their effort levels up through the assessment, and to review their work at the end, thus maximising the formative benefit.

The benefits of smart technology

The technology used to produce on-screen assessments is important – it affects the speed and cost of production, the maintainability and future proofing of the software, and the ease with which the core software components can be adapted for reuse in new assessments.

The approaches used in our tools are based on interchangeable core components using Service Oriented Architecture (SOA) models which provide a range of flexibilities such as:

- Customisable user interface to incorporate client brands and style guides, or compatibility with related tools (e.g. within a VLE).

- Deployment on a range of platforms – CD-ROM, download, web-based, SCORM – without additional work, and with the assurance that the assessment experience is identical, irrespective of factors such as screen size, operating system, etc.
- Inclusion of a range of accessibility and personalisation features as standard.
- Separation of content from platform – allowing items to be edited/cloned/ contextualised without changing underlying software.
- Specialist assessment content management and production platform, allowing multi-role collaboration in content production, and template-based production of powerful item types.
- Opportunity to use a central item bank to deliver on-demand, dynamically created initial assessments and to maintain/update statistical data about each item in the bank.

Some of the challenges of adaptive skills assessments

Using interactive, adaptive assessment does present some challenges. The burden of technical testing is greater than for a simple linear assessment (because of the need to test that the algorithm is functioning correctly, and the need to test a larger number of items than would be present in a linear assessment).

Additionally, adaptive assessments require greater field trialling – not only to check that the various pathways through the assessment are leading candidates to correct level assessments, but also that each of the possible paths provides a suitably broad curriculum coverage.

Generally, tutors and learners expect assessments to be transparent. While it is generally not necessary to include information about the adaptive routing within the product, such information has to be provided in the documentation. Adaptive assessments are more complicated to explain than traditional minimum competence (pass/fail) and compensation-based (graded) assessments.

Interactive assessments which include rich media can present access problems for candidates with disabilities. For example, questions that use "drag and drop" mouse movements are difficult for people with poor motor skills who are using a mouse. While rich media improves the assessment experience for many learners, it must not result in assessments being unreadable by those with poor sight.

Almost every type of media richness and interactivity can be made accessible, but this introduces cost and complexity to the development process and must be planned in from the outset.

Products built with this technology

BTL Group Ltd and Alpha*Plus* Consultancy Ltd have many years' experience in developing on-screen and on-paper Skills Check and Initial Assessment tools. The easy-to-administer tools allow the tutor to quickly determine the learner's requirements, while the tried and tested algorithm provides an accurate measure of the learner's skills.

The Skills Check tools and Initial Assessments that BTL and Alpha*Plus* have developed include those for literacy, numeracy and language (ESOL). These tools have been used by a variety of clients including DfES, New Zealand Ministry of Education, Job Centre Plus, SkillsActive, Probation Services, Road Passenger Transport Authority, Skills for Health, Asset Skills, Skillsmart and Improve (Food and Drinks Skills Sector Council).

People we've developed Initial Assessment/Skills Check tools for:

SQA
QCA
cea Rewarding Learning
SkillsActive Shaping Skills for the Future
Llywodraeth Cynulliad Cymru Welsh Assembly Government
DWP Department for Work and Pensions

Programmes we've developed Initial Assessment/Skills Check tools for:

- Core Skills
- Essential Skills
- Key Stage 3 SATS
- Key Skills
- Skills for Life
- Literacy, Numeracy and ICT in the National Curriculum

Alpha*Plus* Consultancy Ltd

www.alphaplusconsultancy.co.uk

Alpha*Plus* is a team of highly experienced educational consultants with a large and varied portfolio of work in education and training. The team's expertise lies in vocational and general qualification development, assessment and public examinations, e-Assessment, e-Learning and e-Portfolios. Alpha*Plus*' work to date includes high profile and high volume e-Assessments in the UK: Key Skills, Skills for Life, the Driving Theory Test, and the Life in the UK citizenship tests. More recently, Alpha*Plus* has undertaken e-Assessment development work in Functional Skills.

BTL Group Ltd

www.btl.com

BTL is a leading UK assessment and learning development and delivery company. With over one million certified exams powered by BTL, the company's content development and delivery tools are helping some of the UK's largest agencies face the challenge of on-screen delivery. Since 1985, BTL has been successfully helping its customers exceed their assessment and learning expectations, and developing some of the most innovative and engaging on-screen content in the world.

This white paper is published jointly by Alpha*Plus* Consultancy Ltd and BTL Group Ltd. We have worked together in e-Assessment strategy, development and deployment since 2000.

BTL Group Ltd, Salts Wharf, Ashley Lane, Shipley, West Yorkshire, BD17 7DB.