The Design of LearnFlex Evaluator™: a web-based adaptable assessment and evaluation application

Gary Woodill, Ed.D.
Chief Learning Officer
Operitel Corporation

Michael Skinner
Chief Executive Officer
Operitel Corporation

Authentic Assessment

We can't look inside a person's head to see what they know or how much they have learned. Instead, we use indirect measures of knowledge and learning based on observations of what a person says and does. Based on these observations we make judgments about the level of a person's knowledge or progress in learning. This process is formally known as "assessment and evaluation".

Traditionally, assessment and evaluation has been the task of a teacher who has direct contact with a learner. Over the years, various assessment tools have evolved to assist in the systematic gathering of data from observations of learner behaviour. These tools include examinations and tests, observation checklists, rating scales, and surveys. Exams and tests usually have a variety of question types ranging from fill in the blank, true and false, short written answers, and multiple-choice.

Traditional assessment is seen by many educators as inadequate in that a small amount of data is used to describe and summarize a person's total learning and knowledge. There is now a movement in education towards "authentic assessment", an effort to have a more holistic approach to assessment and evaluation. Authentic assessment is a move away from traditional multiple-choice, standardized achievement tests to the inclusion of direct observations of learner performance on meaningful tasks that are relevant to the learner. Burke (1999) lists the following descriptions of authentic assessment:

- Methods that emphasize learning and thinking, especially higher-order thinking skills such as problem-solving strategies
- Tasks that focus on students' ability to produce a quality product or performance
- Disciplined inquiry that integrates and produces knowledge, rather than reproduces fragments of information and others have discovered
- Meaningful tasks at which students should learn to excel
- Challenges that require knowledge in good use and good judgment
- A new type of positive interaction between the assessor and the assessee
- An examination of differences between trivial school tasks (e.g., giving definitions of biological terms) and more meaningful performance in non-school settings (e.g., completing a field survey of wildlife)
- Involvement that demystifies tasks and standards (pp. xx-xxi)

A move towards authentic assessment does not mean discarding traditional assessment tools. Rather, the goal is to strive for a "balanced assessment" that uses traditional testing methods as well as newer methods such as evaluation of portfolios or direct observations of performance.

Measured against the authentic assessment movement, currently available online assessment packages are woefully inadequate. Most consist of tests constructed from 5 to 10 question types, with data being reported as a global score or a percentage of "correct answers" on the

test. Almost all tests are modelled on traditional “paper and pencil tests” that are easy to score.

Almost no online assessment programs focus on the reasons for “wrong answers” to questions. Yet there is value in knowing which wrong answer a candidate has selected on a test. This information can be used for generating adaptive feedback, for setting the direction of further testing, and for diagnostic purposes. And, because answers are often not entirely wrong or right, there is the possibility of giving partial credit for answers that are partially correct but which are not the best or optimal answer.

Before designing LearnFlex Evaluator, an extensive review and “mind map” was developed of the assessment and evaluation literature, online assessment packages currently being offered on the Internet, and technologies being used to support workplace performance assessment and support. From this review, a conceptual model of computer-assisted assessment (CAA) and computer-based assessment (CBA) was created (Figure 1). LearnFlex Evaluator is designed as a specific software application based on this generic conceptual model.

LearnFlex Evaluator is the first online assessment tool that supports a balanced approach - mixing test data with performance assessment data to achieve the goal of authentic assessment. Because there are many aspects of the assessment of human learning and knowledge that cannot be assessed directly by computers, LearnFlex Evaluator allows the manual input of assessment and evaluation data, which can then be combined with data gathered directly by LearnFlex Evaluator or by third-party assessment software.

The Uses of Assessment Results

Most online assessment tools don’t take into account the uses of the assessment data that they generate. Seven different uses of assessment results can be identified:

- *Formative Assessment and Evaluation* - used for feedback while learning (e.g., “continuous improvement” methods)
- *Adaptive Assessment and Evaluation* - used to change or individualize the curriculum or further assessments
- *Demonstrative Assessment and Evaluation* - used to show what a person has done and can do, often compared to a specific set of criteria
- *Summative Assessment and Evaluation* - summarizes achievement levels, assesses program effectiveness, and is used as criteria for granting credentials
- *Diagnostic Assessment and Evaluation* - used to diagnose the presence of a defined diagnostic construct
- *Normative Assessment and Evaluation* – summarizes the results of assessments from many learners to give a set of statistics (or “norms”) for the identified group
- *Research* - assessment is used to answer specific questions, make predictions or supply data for other software applications

The current version of LearnFlex Evaluator supports the first four uses of assessment listed above, and can be customized by Operitel to support the gathering of data for specific research questions or for interacting with other software. Future versions of Evaluator will incorporate diagnostic and normative aspects of assessment.

An extensive search of the assessment literature was used to build a large "mind map" of the entire field of assessment (available from the first author as a Visio file). From that exercise, a model was developed of online assessment processes. Figure 1 shows a conceptual model of computer assisted assessment (CAA) and computer based assessment (CBA). The main difference between CAA and CBA is that in CAA the computer assists the assessor by providing resources, while in CBA the computer carries out the entire assessment and reports its results.

Assessment Constructs

What initially comes to mind in talking about assessment and evaluation is the testing of the ability to reproduce information – or "how much we know". This is the traditional view of assessment in schools, where the ability to recall "facts" is of paramount importance. This goes hand-in-hand with the dominant method of teaching in schools, verbal presentations and explication of "course materials" by a teacher at the head of the class, coupled with extensive reading and writing as "homework".

Because of the prevalence of this "Tell-Test" model of education, it is not surprising that most online assessment packages are based on the testing of information recall. While recall may be important in many situations, it is only one aspect of human performance that can be assessed. Howard Gardner (1993), with his concept of "multiple intelligences", has long advocated for a much broader view of teaching and assessment. Therefore, authentic assessment needs to take into account other aspects of learning, such as:

- Learning processes
- Learning transfer and application
- Meta-cognition ("thinking about thinking")
- Creative abilities
- Values and meaning
- Direction and growth
- Physical abilities
- Reasoning abilities
- Collaborative and social abilities
- Leadership qualities

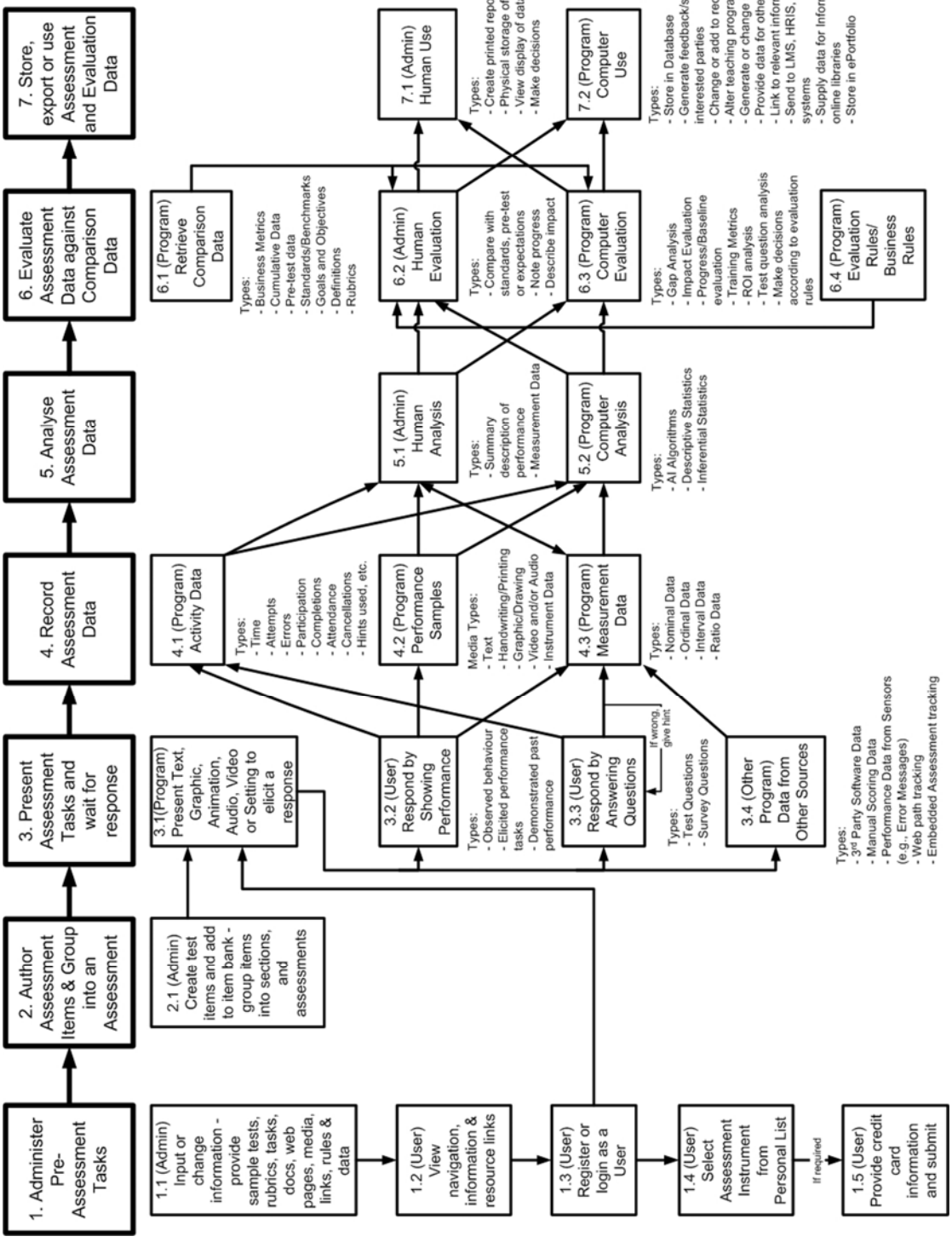

Figure 1- Computer Assisted Assessment (CAA) and Computer Based Assessment (CBA) – a conceptual model

As well, assessment often needs to go beyond the simple reporting of numbers to the identification of various educational, psychological, sociological and neuro-physiological constructs. Assessment constructs include such ideas as IQ, personality traits, medical and psychological diagnoses, performance levels, and attainment of specific competencies. These constructs are often identified by specific answers, or a particular combination of answers, to test questions or performance indicators. Future versions of LearnFlex Evaluator will support the identification of specific assessment constructs through the tagging of specific questions, and/or answers to questions, and its ability to allow assessors to write simple rules to identify such constructs.

LearnFlex Evaluator is a powerful and adaptable assessment engine that allows a wide variety of assessment protocols. It features several of supporting performance assessment online and over 20 types of online test questions. When linked to the LearnFlex learning management system (LMS), it allows data from third-party software or from manual scoring to be easily entered into the LearnFlex Evaluator database and used with the performance data and/or test results that it has collected. In the background, LearnFlex Evaluator can be collecting activity data such as times, attempts, participation, attendance, and errors. This can support the assessment picture being developed, and add to its completeness. The LearnFlex Evaluator database can hold a wide variety of comparison data such as previous scores, standards and benchmarks, and assessment rubrics. This allows for gap analysis, needs assessments, and other learning metrics. Integrated with a learning management system or a human resources information system, LearnFlex Evaluator can feed information to many other enterprise applications.

LearnFlex Evaluator: supporting the entire assessment and evaluation process

LearnFlex Evaluator supports the entire assessment and evaluation process from beginning to end. It is an online assessment resource library, an assessment authoring system, an assessment delivery engine, an assessment evaluation engine, and an integrator and manager of assessment data from many sources. LearnFlex Evaluator supports assessment and evaluation even if a learner is not taking an assessment online. In its Computer Assisted Assessment (CAA) role, online library for LearnFlex Evaluator is able to store and print a wide variety of tests, descriptions of performance tasks, scoring rubrics, reporting templates and assessment documentation. These are available on request for authorized personnel from the Assessment Library. Data from the results of manually scored tests and performance tasks can later be entered into the LearnFlex Evaluator database and be further analyzed and distributed.

Computer-Based Assessment (CBA) with LearnFlex Evaluator comprises seven distinct stages:

1. *Pre-assessment Administrative Tasks*: LearnFlex Evaluator has extensive administrative functions that need to be used or reviewed before it can be used for assessment and evaluation. Administrative tasks include the configuration of the user registration system, the uploading of media clips to be used in an assessment, the construction of introductory web pages, and the provision of appropriate feedback and/or hints for each question or task. Scoring rubrics, assessment documentation, and links to other assessment sites can be entered into the Assessment Library at any time. System properties that govern such aspects as the "look and feel" of the navigation, translations into other languages, and evaluation and business rules used to process results can all be added to LearnFlex Evaluator.

2. *Assessment Authoring*: in this stage, tests, test questions and performance tasks need to be constructed in LearnFlex Evaluator for later use. Along with the question or task description needs to be all the necessary information to compose, render, score, and provide feedback. The IMS QTI specification defines an "item" as "the Question, layout rendering information, the associated response processing information, and the corresponding hints, solutions and feedback". This level of complexity is supported in Evaluator.

3. *Presentation of Assessment Tasks*: here an assessment task is presented to the user in order to evoke a response that generates relevant assessment data. Broadly speaking, there are four kinds of assessment tasks:
    a) naturalistic observations
    b) solicitation of third party opinions
    c) requested performance tasks
    d) answers to questions

    Naturalistic observations are assessment tasks that are undertaken directly by an observer without the person being observed being prompted to perform a task in a specific way. Solicitation of third-party opinions are techniques that ask others to remember their observations of the person being assessed, sometimes referred to as "360° evaluation". Requested performance tasks involve the person being assessed following a set of instructions from the assessor, and submitting a record of the performance. Answers to questions are a person's responses to oral or written questions.

    LearnFlex Evaluator supports performance assessment in five different ways:

    a) Assessment Resources
       Examples: LearnFlex Evaluator Assessment Library can supply printed tests, assessment documentation, report templates and scoring rubrics on request.
    b) Online performance with computerized tracking
       Examples: interaction with a simulation or a game that gives a score, elapsed time on solving a problem, or any other online activity that can generate AICC or SCORM data.

c) Submission of a performance sample in digital format. Users are able to upload such samples to a third party for evaluation.
d) Links to an e-portfolio
   Examples: can be an object repository in MS-SharePoint, or a personal library in an LMS or in LearnFlex Evaluator.
e) Online tools for real-time performance observation
   Examples: self-scoring online rating scales or checklists.

Online tools that can be used for direct or indirect performance assessment (for self or others) include:

- Checklists with radio buttons (one choice per item)
- Checklists with check buttons (multiple choices per item)
- Ranking scales
- Rating scales
- Surveys/questionnaires/polls
- Online notebook

Tests are a special type of performance task in that the person being observed is asked to answer a variety of questions. Components of tests are Sections, Tasks ("subsections"), Questions, and Answers.

There are a wide variety of test questions that can be used in online assessments. These include:

- Checklists
- Calculation Questions
- Classification Questions
- Interpretive exercises
- Multiple Choice Questions
- Multiple Answer Questions
- Locate Answer Questions
- True-False Questions
- Masked Answer Completion Questions
- Numeric Answer Completion Questions
- Drop-Down Completion Questions
- Extended Response Essay Questions
- Restricted Response Essay Questions
- Matching Questions
- Drag and Drop Questions
- Likert Scale Questions
- Rank Order/Sequencing Questions
- Quick Reaction Questions
- Reconstitute Sentence Questions
- Crossword Puzzle Questions
- Multi-part Questions

4. *Record Assessment Data*:

   Assessment data consists of the results of automatically scored online assessments, records of related activities while the assessment items were being answered, or are the results of other assessments where the results are manually entered, or are assessment data received from another software application.

5. *Analyse Assessment Data*:

   Once assessment results have been recorded and stored, they can be analysed through various statistical procedures (using third party software), and through human analysis. The results of human analysis can then be manually entered into LearnFlex Evaluator for further processing.

6. *Evaluate Assessment Data against Comparison Data*:

   A powerful feature of LearnFlex Evaluator is its ability to compare assessment results with other stored data, and to render a decision (e.g., “pass” or “fail”) based on this comparison and a set of “evaluation rules”.

7. *Generate Reports, or Store, Export and Use Assessment and Evaluation Data*:

   Assessment and evaluation data can then be used by LearnFlex Evaluator in several ways, including:

   - Data can be stored as a cumulative record for an individual
   - Printed reports can be produced on individual assessment and evaluation results
   - Aggregate assessment and evaluation data can be put into group reports
   - Aggregate assessment and evaluation data can be supplied to an executive “dashboard” (in the LearnFlex LMS)
   - Feedback messages can be sent to the assessment candidate, and other interested parties
   - Cumulative records and group statistics can be changed in regards to a specific assessment
   - Data can be exported to a learning management system so it can be used to complete competency lists, and then modify the presentation of educational materials
   - Data can be exported to another enterprise application such as a HR Information System or a Student Information System
   - Data can be exported to a MS-Excel spreadsheet

Assessment Issues

The use of any online assessment tool does not solve all issues in the field of assessment. There is an extensive literature on issues with the methods, interpretation and use of assessment and evaluation data. Issues include:

- Lack of assessment training for teachers
- Teaching to the test
- Effects of teacher expectations
- Overuse of standardized testing
- Assessing unique constructions of knowledge
- Assessing "whole task" performance
- Assessing values
- Assessing collaborative learning
- Assessing non-linear and continuous learning
- Assessing "multiple intelligences"
- Assessing learning processes
- Variations in grading practices

While many issues will remain problematic, LearnFlex Evaluator can help to deal with many of the problems listed above. Lack of teacher training can be alleviated by an online library of assessment templates and rubrics. Teaching to the test and scoring biases due to teacher expectations can be countered with random selection of items from a comprehensive question bank. The ability to assess learner performance in several different ways allows for a more balanced and holistic approach to assessment and evaluation. Sampling of learner interactions and activity during learning activities (for example, working through a tracked interactive simulation) or differential analysis of test items and/or sections can produce indicators of learning processes and diagnostic categories.

LearnFlex Evaluator is an innovative, full-featured assessment and evaluation application that produces detailed information for multiple uses. For more information, contact Operitel Corporation at 705-745-6605, or e-mail Gary Woodill at gwoodill@operitel.com.

References:

Burke, Kay (1999) *How to Assess Authentic Learning.* 3rd Edition. Arlington Heights, IL: Skylight Training and Publishing.

Gardner, Howard (1993) *Frames of Mind: the theory of multiple intelligences*. New York: Basic Books.