Journal of Universal Computer Science, vol. 4, no. 4 (1998), 330-348
submitted: 10/12/97, accepted: 28/12/97, appeared: 28/4/98 © Springer Pub. Co.

Assessment in WWW-Based Learning Systems: Opportunities and Challenges

Mary Hopper, Ph.D.
(Networked Multimedia Information Systems (NMIS),
Massachusetts Institute of Technology, Cambridge, MA, USA
mehopper@ceci.mit.edu)

Abstract: This paper proposes that the most important role the computer may play in education could be contributing to the ubiquitous use of assessment for the improvement of instruction. In order to realize this potential, newly emerging WWW-based learning systems should support a very wide range of embedded assessment features. These systems require architectures with a core of reliable integrated management tools, one or more modules with instruction and assessment, standard database connectivity, and an acceptable level of attention to permissions and security. No company will adequately address all of the possibilities for assessment in WWW-based learning systems, so it is critical that WWW based learning systems have "open system" architectures and company policies for cooperating with other companies to support interoperable modules. The point is raised that some products with similar types of assessment features can have very different architectures and policies for supporting interoperable modules. It is recommended that "checklists" for comparing assessment capabilities should be viewed with skepticism, because they can favor products with weaker architectures and policies for accommodating assessment capabilities.

Category: K.3.1
Key Words: educational assessment, electronic learning systems

1 Introduction

When computers are employed to perform tasks traditionally performed by humans, the question to ask is "what is to be gained?" In the area of educational assessment, the question becomes "what does the computer contribute to the assessment process that would otherwise be more difficult or impossible?" Unfortunately, this question is complex and the answer is not straightforward. The practice and theory of educational assessment is best understood as on-going attempts to define its role and methods, both of which are still in process [Scriven 94]. The situation is no better within the specific area of educational computing [Reeves 92]. The widespread lack of agreement is undoubtedly part of the reason that evaluation of educational computing innovations has seldom been done outside of demonstration projects. While there are initiatives designed to improve the situation [EvNet 96] [Hawkes 97], the problems remain hard to resolve because of the wide range of views [Anderson and Draper 91].

2 A Time of Opportunity

Despite difficulties that will be hard to overcome, recent developments make this a critical period to define how computers can enhance the assessment process. There is a trend across educational institutions to create World Wide Web (WWW) oriented networking infrastructures to provide information and integrated computing services directly to faculty, staff, and students [Oblinger 95]. Such systems could evolve into defacto educational computing systems as their capabilities expand into campus-wide course delivery and assessment.

There is also a surge of products for creating, delivering, and managing instruction on WWW-based systems. Wide spread use of these products is being facilitated by the emergence of widely accepted standards [McCollum 97]. One leader in the area of industrial courseware development issues, the Aviation Industry Courseware Committee (AICC), recently released a standard to ensure the interoperability of computer-managed-instruction (CMI) and computer-based-training (CBT) products from different vendors [AICC 97]. The list of companies supporting the standard is large and includes Apple, traditional integrated learning system companies WICAT and TRO, WBT Systems (maker of TopClass), and Macromedia (maker of Authorware). The membership of the alliance, combined with the publicity that surrounded the announcement of this standard, signals a serious interest of key companies in this area.

The AICC standard is not going to influence all computer-based learning systems immediately, but it’s success suggests the maturity of a market necessary to support standardization. If trends in other software areas are an accurate indicator, a few major companies will soon overwhelm most competitors and usher in an era of fewer packages with wider use and availability. This could bring the long predicted wide spread use of computers in education, if textbook publisher’s choose to deliver content via these products. This speculation is loosely supported by the closely timed alliance of McGraw Hill Educational and Professional Publishing Group with two AICC supporters, CBT System (TopClass) and Macromedia (Authorware), who are also top contenders in the market of WWW-based learning systems [Odell 97].

A standardization of WWW-based learning systems will result in either gains or loses for assessment practices. The assessment functions associated with textbook-based instructional programs will be transferred to the computer, and then constrained or expanded based on the assessment features included in widely available products. The characteristics of the products will help to facilitate or deter ubiquitous emphasis on assessment, and also influence the forms and quality of assessment that are implemented. If limited capabilities are included in the emerging standard packages, opportunity will be lost. As critical products are developed, it is essential that companies are shown the potential for what can be done, and receive pressure from the educational community to expend the resources to do it. If this is not done now, it will be far harder to correct the situation later after the products become widely adopted.

The remainder of this paper will focus on the following tasks:

- Describe a variety of potential approaches for using computers to augment the educational assessment process.

- Determine the key characteristics required to support those approaches in WWW-based learning systems.
- Compare leading products relative to the characteristics required to support a broad range of assessment practices.

3 Assessment Roles in Computer-Based Learning Systems

Rather than try to take on the daunting task of definitively addressing "THE ROLE" of the computer in the assessment process, this paper will focus on describing a few ways the computer can augment a range of conceptions of assessment in education. The following section will describe a sample of the potential advantages the computer brings to the assessment process, framed within a historical review of how those advantages have been realized in various forms of computer-based learning systems.

3.1 Classical Role of the Computer in Assessment: Measurement Manager in "Integrated Learning Systems"

One obvious way that the computer can improve the assessment process is at points where number collection and manipulation are needed. Thus, the initial role of the computer in the educational assessment process was as a capable partner for processing numerical measurement data collected primarily with objective tests for use in the evaluation of learners. By 1971, Thorndike could describe a history of this "classical role of the computer" for augmenting traditional measurement processes through "computerized testing." Thorndike pointed out that the computer could easily generate, store, administer, collect, analyze and report assessments automatically [Thorndike 71].

Twenty years later, this view of the computer's role in assessment was only moderately expanded by Stager and Mueller in the article "Computer Uses in Classroom Testing" to include the ability to automatically provide detailed feedback to students about their strengths and weaknesses. However, they note that one of the reasons that the computer was not used more by classroom teachers was that the way it was used did not encourage integration with instruction [Stager and Mueller 91]. By addressing this weakness, the computer evolved to the “classical role” of managing the measurement process for evaluating learners within Integrated Learning Systems (ILS). Instruction and assessment were integrated in ILS products through an emphasis on using objectives to align both.

At first glance, the emerging WWW-based learning systems appear to be re-implementations of ILS products. Reviews of the new products make claims like "many different types of assessments," "questionnaires and variety of assessments," "multiple choice assessments and feedback" (taken from actual reviews of popular packages). Caution is warranted in interpreting such claims. Closer examination can reveal that some products lack much of the sophistication of traditional ILS products. As outlined in Table 1 below, some products only include limited features which automate the tasks that educators already perform, while omitting extended features that provide improvements to the assessment process that are impractical to achieve without the computer (<see> Tab. 1).

Table 1: Classical Role of the Computer in Assessment

Classical Roles	**Limited Features**	**Extended Features**
Item Bank	• Packaged Pre-Constructed Item Sets • Expandable Item Entry and Storage • Item Type Multiple Choice • Selection Questions (e.g. matching, yes/no, true/false) questionnaire/Surveys (Likert-type rating scales) • Completion, Short Answers, Essay (no processing)	• Multiple Classifications (Objective, Level, Type) • Graphical Structuring Tools (Outline, Table of Specification) • Classical Item Analysis • Response History (Difficulty, Discrimination, etc.) • Distracter Evaluation and Flagging • Multiple Response (none, one or more alternatives) • Free-Response Evaluation (3,000 characters) • Essay (Unlimited Length) with On-Line Marking • Complex Problem Sets (Numbers, Scripting and Multimedia Features)
Test Generation and Admin-istration	• Editing / Layout / Printing • On-line Administration • Timed, Interruptable, Error Trapping	• Equivalent Forms • Multiple Versions Generation • Multiple Scales or Subtests • Item Sampling • Adaptive Delivery
Test Scoring, Analysis, Record Keeping (Gradebook)	• Data Entered by Class and Name • Calculation of Cumulative Scores • Group and Individual Reports (Total Score Only)	• Major Types of Statistics (Mean, Median, Mode, Range, Standard Deviation and Error, KR21 Reliability, Z / T Scores etc.) • Individual Item Performance Report • Individual Diagnosis and Prescription Analysis • Norm-Referenced and Criterion Referenced Analysis and Reporting • Help Facilities for Score Interpret.
Student Feedback	• Raw Score • Class Distribution and Mean	• Detailed Item Performance Analysis • Detailed Diagnosis of Individual Strengths and Weaknesses • Individualized Task Assignment

If WWW-based learning systems simply mirror the pragmatic weaknesses in the assessment process practiced by educators in traditional settings, there will be a lack of emphasis on special advantages that the computer can provide. For example, products may include multiple-choice tests, but lack a feature to generate classical item analysis data for improving test items. This is a valuable practice that has been impractical in day-to-day classrooms because of unrealistic expertise, time, and computational requirements. Many products also lack structured approaches for creating and managing instruction and assessment with detailed objectives, yet this was the key feature which brought about the improved integration of instruction and assessment found within traditional ILS systems.

WWW-based learning systems are relatively new and under development, so product release notes may contain valid claims that there will be more extended features in future releases. However, some widely accepted products may continue to lack valuable extended assessment features because these features have more data intensive requirements. For example, some of the most valuable extended features virtually require an architecture that can include standard databases for tracking learner performance (ODBC and JDBC connectivity as opposed to relying on flat files) [Dynes, S., Litchfield, L., Curtis, K. and Chiquito, A. 97]. Unfortunately, it is possible that products will remain competitive with limited database connectivity by omitting data intensive assessment features. This would result in marketable products that do not provide the features needed to realize the potential improvements in the assessment process that the computer can offer.

3.2 Innovative Role of the Computer in Assessment: Monitor in "Interactive Learning Systems"

The most important role that the computer can play in the assessment process is not the "extended features" described above. A powerful force behind computer use in education has been predictions that innovative uses could lead to improvements in learning. Claims have ranged from more reliable and valid ways to assess high-order thinking skills to producing revolutionary impacts on human cognitive abilities. It is an understatement to point out that the full range of these claims can not be adequately addressed on this page or even in a single paper. However, products based on these claims, loosely referred to as "Interactive Learning Systems," share the common focus of providing improvements to both instruction and assessment beyond those provided by traditional "Integrated Learning Systems." Proponents of innovative uses of computers tend to agree that traditional computerized testing inappropriately equates "evaluation" to "testing" (esp. multiple choice). Proponents of "Interactive Learning Systems” recognize that measurement legitimately includes a variety of methods. These innovative approaches use a variety of measures for making evaluations, and either demote traditional tests to simply another instrument for collecting data or omit them altogether [Knussen, Tanner and Kibby 91]. The following are three rough groupings of innovative approaches to using computers in education with corresponding examples of approaches to assessment that may eventually be embedded within learning systems.

3.2.1 Exploration and Interaction with Constructed Experiences

There are a few different types of Interactive Learning Systems that emphasize exploration and interaction. One popular approach uses interactive graphics for achieving higher-order objectives in problem-sets. Another approach is creating intelligent microworlds that allow learner's to interact with pre-constructed simulations which use artificial intelligence (agents) to construct a model of the learner's misconceptions for use in diagnosis and prescription of on-going activity [Feurzeig 87]. Hypermedia environments can range from simple exploration and interaction with navigational maps, tours, and indexes [Hammond and Allinson 89] to complex multimedia simulations (including virtual realities) [Henderson 91]. These approaches all share a focus on assessment strategies based on automatically capturing and analyzing data about user's actions. These approaches are also more likely to place an emphasis on using assessment strategies for less traditional concerns like learning style, metacognitive strategy, and motivation [Reeves 92].

3.2.2 Generative Learning with Focus on Learner Construction

Interactive Learning Systems that emphasize generative learning can include the most innovative approaches to both instruction and assessment. While problem-solving is still emphasized, the focus is on having learners to construct their own knowledge representations rather than interacting with those created by others. These approaches are often associated with Papert's "Constructionist" pedagogy. Thus, they include construction of mathematical models with computer languages such as Logo [Papert 80] and multimodal representations with hypermedia [Harel and Papert 90]. Some groups focus on specific contexts such as representing literary interpretation by students using conceptual mapping tools [Landow 89] [Yankelovich, Meyerowitz, and VanDam 85].

Generative approaches share a concern for assessing learner constructed products and complex performances of demonstrated capability in contexts. Evaluation in these approaches is often based on efforts to directly determine the type, amount, and quality of contributions to either an individual portfolio or group corpus. There is frequently a focus on performance-based assessment. Evaluation may be based on observations of progress interviews, demonstrations or presentations, and may also turn to open-ended ethnographic techniques such as structured observational studies. Videotape is used to capture these for review, and the computer may be used to manage everything from the notes to the actual digitized video [Bennett and Hawkins 94]. Proponents generally believe that traditional forms of “objective-based” measurement are unlikely to uncover true innovation, so they use open-ended methods that exploit opportunities to capture serendipitous developments and unanticipated emerging outcomes [Baker, Herman and Gearhart 96].

3.2.3 Cooperative and Collaborative Learning

In addition to using the same assessment techniques as the other innovative approaches, educators that emphasize cooperative and collaborative learning often focus on examining participation in learner-to-learner interactions in electronic communication (E-Mail, Discussion Forums, Newsgroups, Chat and Multi-user Environments such as MUDs or MOOs, and Groupware). There is a focus on automatically recording and monitoring such variables as the amount, length, and type of on-line interaction. This includes data such as the total number of contributions, total hours on-line, total number of log-ins, total number of messages sent, length and quality of interaction, and finally interaction patterns analysis based on number of communication partners [Hiltz 90]. More sophisticated approaches to communication patterns analysis can include "intermessage reference analysis and maps," "message act analysis" (initiations or replies), and "message flow analysis and diagrams" (density of message over time) [Levin, Kim and Reil 90]. In the near future, these approaches will also undoubtedly adapt these strategies to contributions via white board, tele-conference, and video-conference.

3.3 Emerging Role of the Computer in Assessment: Accommodating Approaches within "Open Learning Systems"

The above are a few examples of innovative uses of computer-based learning systems and some corresponding implications for assessment. The intention has been to demonstrate a key point about the potential of computers to augment the educational assessment process, rather than to present an exhaustive list. While the extended features of "classical" approaches depend on the ability to use standard databases, more innovative approaches place far greater demands on the architectures of WWW-based learning systems. Some packages may be extended to include capabilities for a few of the most popular innovative approaches to instruction and evaluation, but no single product is likely to accommodate them all adequately. For this reason, innovative approaches to instruction and assessment require the capability to extend WWW-based learning systems with products from multiple vendors. This is what is needed so that ambitious authors can work with small vendors to add exciting new features that accommodate new or different approaches to education.

It may be argued that the best approach to providing a variety of instruction and assessment capabilities ranging from "classical" to "innovative" is to simply use different packages. Such an endeavor could prove inappropriate and short-sighted. WWW-based learning systems will soon become both standardized and integrated with broader campus-wide information infrastructures. If innovative packages are supported piece-meal, they will stay removed from the mainstream and be considered "hassles" because of poor integration. The most innovative uses of computers would thus remain the province of a few dedicated instructors who demand to use them despite their inconvenience. To accept this would result in a missed opportunity. This point can be better understood through a closer examination of the contradictions that are often assumed to exist between "classical" and "innovative" approaches.

While the common interpretation is that innovative approaches represent a predominant concern for learners, a more accurate representation is found within Lawler's portrayal of the central dilemma of education as a concern for "instructing" while respecting "the self-constructive" character of the mind [Lawler 82]. A deep concern for "instruction" is not absent from this formulation. In fact, "innovative" approaches are based on the belief that in order to communicate instruction effectively, it must be tailored to the nature of the learners. Educators dedicated to communicating a discipline can eventually realize that attempts to "instruct" without taking the nature of the learner into consideration tend to be less effective. Thus, instructors evolve from using the less "learner oriented" classical approaches to the more "learner oriented" innovative approaches [Hopper 93].

The use of the computer in education may evolve from traditional and familiar to innovative and unique, so the role of WWW-based learning systems should be accommodating diverse approaches to instruction and assessment, rather than dictating or constraining them. The architectures of computer-based learning systems should be "open" enough to support the widest range of possibilities. If they are not, one of the following two detrimental situations might occur:

• If systems only accommodate "classical" approaches, then instructors will be constrained to the traditional uses that they feel comfortable with at first, rather than the more innovative approaches to which they might gradually evolve towards.

• If systems only accommodate "innovative" approaches, then instructors will not be ready to appreciate the innovations, even though they might have eventually grown to appreciate the system's features over time in they had a chance to adapt gradually.

This is why the architectures of learning systems should be capable of supporting a variety of modules with tightly coupled and congruent instruction and assessment components. If the packages that become standardized are not such "open learning systems", then in the long run it could lead to lost opportunity for expanding the use of the most innovative opportunities offered by the computer for improving both instruction and assessment.

3.4 Expanding Role of the Computer in Assessment: Continuous-Improvement of "Open Learning Systems"

The most important role of the computer in education might ultimately be contributing to widespread feedback for the improvement of the quality of instruction, rather than simply serving as a better tool for delivering instruction and assessments to learners. The key to realizing this valuable possibility is by accommodating an expanded view of assessment within WWW-based learning systems.

3.4.1 Assessment Focused on Improving Instruction

While the assessment of learners is one potential improvement to the assessment process offered within computer-based learning environments, there are a variety of other roles the computer can play. In fact, differences in assessment based on approaches to instruction are seldom the main distinctions considered in either discussions of assessment in general, or computer-based learning systems in particular. Discussions about educational assessment can, and usually do, refer to making determinations about the quality of the instruction, rather than the learners. While both share the goal of improving instruction, the assessment of instruction focuses on top-down determinations of quality and accountability rather than bottom up improvements within learners. Sometimes it is hard to tell the difference between the two, and they are not mutually exclusive. The same strategies can be used in both processes, and much of instructional assessment can be embedded in regular instructional activities. One traditional way of framing this distinction is based on the form of evidence considered in the assessment of instruction. In Stufflebeam's famous CIPP model, learner performance is considered "Product," while the other considerations of assessment are "Context," "Input," "Process" [Stufflebeam 73].

When assessment focuses on aspects of the instruction rather than the learner, then specific elements of the instructional situation are considered individually. Thus, assessment can include determinations regarding the quality of the instructor, instructional technology (method), and information technology (media). The distinction between assessment of the "information technology used to deliver" from "instruction delivered" are particularly important. "Information technology" includes the equipment, machines, and media that provide access to instruction, while "instructional technology" refers to strategies for instruction such as the sequence and structure of lessons, the use of examples, or provisions for practice [Clark 94]. Of course, different approaches to assessment of the learner also emphasize different approaches to the assessment of instruction.

3.4.2 Assessment for Supporting a Cycle of Continuous Improvement

Another traditional distinction in assessment has generally been between formative and summative evaluation. Formative information is used as feedback to help improve instruction, while summative information is collected at the end of instruction to be used by decision makers in determining the value of instruction. The most salient issues for this distinction are timing and audience (i.e. during instruction for improvement by participants or after instruction for decision making by stakeholders). However, distinctions based on the timing of assessment can be somewhat artificial. Instruction in distributed computing environments like the WWW can be so dynamic, it becomes difficult to identify a specific point at which to gather "summative data." Many approaches to assessment have moved away from an emphasis on tests for feedback about whether an outcome was achieved, to unobtrusive continuous-measurements that provide feedback on the process of obtaining outcomes [Flagg 90]. Under these circumstances, assessment is used within a cycle of continuous improvement. The sampling of performance is replaced with the monitoring of a dynamic system in which there is a steady flow of information. In this approach, traditional "summative" functions are addressed by taking snapshots of the system to make determinations regarding the quality of its functioning [Newby, Stepich, Lehman and Russell 96].

3.4.3 Assessment for Broader Audiences

A less salient distinction in the conception of formative versus summative information is between an audience of participants (learner and instructor) and an audience or of stake-holders (decision makers, parents, or community members). Of course, some data can be used for multiple audiences. One might expect that embedded assessment features within computer-based learning systems would have been emphasized by stake-holders, since there is the potential for such large unjustified expenditures of resources relative to many other types of educational practices. But this is not the case, and extensive assessment is seldom performed in either traditional contexts or forefront research projects.

It might be especially useful to support institution-wide instructional assessment functions. For instance, WWW-based learning systems could contain an intelligent "Instructional Inventory Generator" to automatically pole and construct instruments to measure instructional variables relative to specific objectives. These could prove useful for situations in which instructors and stakeholders want feedback about the progress and quality of instruction [LeBold, Montgomery and Ward 90] [LeBold, Budny and Ward 97]. It might also be useful to support the automatic administration of institution-wide assessments of instruction with automatically maintained norms for comparison. This would serve the same function as the assessment instruments used by some universities. One example of this sort of assessment is the Cafeteria Course Evaluation Survey used at Purdue University. The Cafeteria is administered to students to measure their opinions about instruction [Allan, Starry and Wright 73] [Derry, Seibert, Starry, VanHorn and Wright 74]. Ironically, these sorts of evaluations are often administered on paper, and then scanned into computers to be processed, and then reported back with paper. The process would certainly be more efficient if done entirely by computer.

The integration of automatic assessment data collection and management on a broad range of instructional variables for a variety of audiences could result in a richer set of available empirical information for use in justifying and improving computer-based learning systems. The following table illustrates a broad array of assessment functions a WWW-based learning system might support automatically, although it is a small sample of the breadth of potentials for both evaluating learners (<see> Tab. 2) and instruction (<see> Tab. 3).

Table 2: Matrix of Learner Assessment in "Open Learning Systems"

	Target Audience		
Focus	Stakeholders	Instructors	Learners
Learners			
Classical Approach	Demographics Student Background Attendance Grades Achievement Tests Questionnaire Data Survey Data Systematic Rating Advancement Record Certification	Student Profile Pretest Prescriptive Test Achievement Tests Grade Calculation	Feedback for Self-Improvement
Innovative Approach	System Activity Self-Assessment Survey Free Response Survey Opinions On-Site Observations Interviews Portfolio Products	Performance Logs Interviews Discussion Ratings Collaboration Ratings Continuous Measures Open-Ended Reports Portfolio Evaluation	Introspection Tools

Table 3: Matrix of Instructional Assessment in "Open Learning Systems"

	Target Audience		
Focus	Stakeholders	Instructors	Learners
Instruction			
Classical Approach	Needs Assessment Cost Benefit Analysis	Questionnaires Surveys	Course Profiles Course Ratings
Innovative Approach	Open-Ended CaseStudy	Observation	Mentor or Peer Ratings
Instructor			
Classical Approach	Instructor Profile Systematic Rating Peer Rating/Comment Expert Reviews Accountability Measure	Open-Ended Surveys Peer Review	Instructor Profiles Instructor Ratings
Innovative Approach	Activity Journals, Logs, Reports Attitude and Belief Inventories Document Review (e.g. lessons)	Self-Performance Journals/Logs Self-Reflection Tools Feedback from Monitors (Responsiveness, Involvement, Facilitation etc.)	Mentor or Peer Ratings Awards Based on Ratings Instructor Style Data
Instructional Technology			
Classical Approach	Implementation Surveys/Reports Time and Training Requirements	Tryout Review/Report/Rating	Peer Reports
Innovative Approach	Instructor Reflections Observations/ Reports	Aggregate Usage Data/Logs Spontaneous Help Message Log FAQs Statistics/Logs	Learner Controlled Selection
Information Technology			
Classical Approach	Implementation Survey Training Requirements Availability/Stability Service Records	Tryout Review/Report/Rating	Solicited Rating and Comments Usability Surveys
Innovative Approach	Acceptance Surveys/Reports	Spontaneous Help Message Log FAQs Statistics/Logs Bug Reports Logs	Preferences Survey

3.4.4 Implementing Assessment of Instruction

Embedding the assessment of instruction into WWW-based learning systems should be straightforward, since the formats of the instruments are often similar to what is already available to use for the assessment of learners. The difficulty is more a matter of making the instruments and results available to a broader range of audiences on a routine basis. This requires the technical capability to support a variety of views or access levels. It also includes a more complex system of grouping, permissions, and file sharing. Of course, there are critical issues of data security to be considered if instructor evaluation records are kept within the same system as learner quiz grades.

4 Architectures for "Open Learning Systems": Approaches and Tradeoffs for Assessment

To ensure that current opportunities are not allowed to pass, WWW-based learning systems should support a wide range of embedded features for enhancing the educational assessment process. These include the following:

- Integrated Processes for Instruction and Assessment
- Extended (Maximum) rather than Limited (Minimum) Assessment Features
- Appropriate Assessment for Classical and Innovative Approaches to Instruction
- Focus on Assessment of both Learner and Instruction (i.e. Instructor, Instructional Technology, and Information Technology)
- Non-Intrusive Continuous Assessment to Contribute to a Cycle of Continuous Improvement
- Integrated Collection and Management for Data for Broad Audiences (Stakeholders, Instructors, Learners)

The above features carry a series of implications for the technical requirements of WWW-based learning systems. Products should include the following characteristics:

- Support for Standard Database Connectivity to Manage Intense Sorting, Storing and Returning of Coded Data
- Open Ended Interface Extendible to Interchangeable Modules of Tightly Coupled Instruction/Assessment Components
- Strong Non-Intrusive and Controllable Permissions and Security

Given the characteristics required to support such a range of features, it is clear that no company will cover the breadth of capabilities possible in a WWW-based learning system. For this reason, it is best to choose a product with an adequate architecture to accommodate a number of other companies. There are a number of small, highly developed products that focus on specific tasks that would make valuable contributions to a larger learning system if they could be easily included.

The following are just three examples of these types of products:

- QuestionMark is a sophisticated and popular "classical" testing package. [Assessment Systems Corporation 97]
- cT is a package with a history that stretches back to the Plato system. It includes sophisticated answer processing and interactive problem set

construction tools. [WorldWired 96]
• Story Space is a product that has provided a popular tool set for constructing hypertext. [Bernstein 88]

Products that find ways to incorporate highly respected software modules from other companies are more likely to accommodate a broad range of instruction and assessment capabilities over time. This indicates the need to adapt an infrequently used approach to ILS design which involves developing an ILS shell capable of incorporating software from many different vendors. Using an ILS shell provides the advantage of accommodating diversity, but results in the challenge of integrating separate packages with system wide functions such as student tracking and reporting. This difficulty in "open learning system" design can be overcome through incorporating a common or core "CMI" module to integrate instruction/assessment modules from other vendors [Maddux and Willis 93]. Given this final requirement, combined with the prior list of characteristics for supporting the broadest possible variety of high quality assessment features, it is now possible to describe a potential architecture for an "open learning system" (<see> Fig. 1).

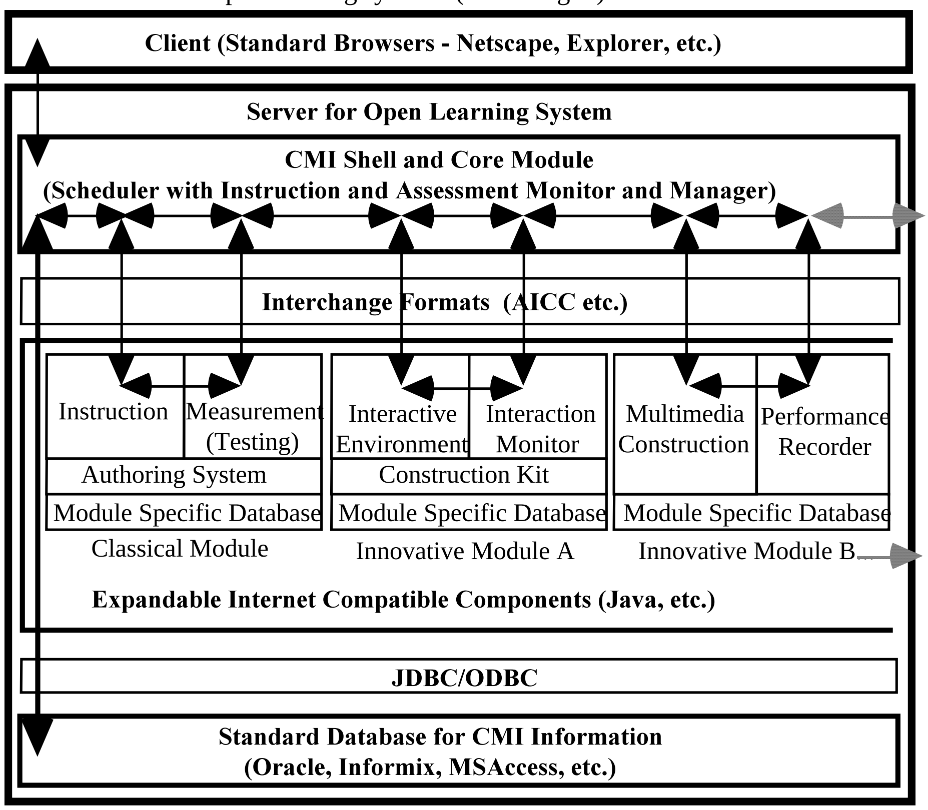

Figure 1: Architecture for "Open Learning System" with Embedded Assessment Features

5 Beyond "Feature Checklists" for Comparing WWW-Based "Open Learning Systems"

Systems constructed with an "open architecture" are more complex, and thus also more difficult to use. Yet usablity is critical, because if systems are not usable, they will not be able to be a required part of a course. If participation is not required, then they are not used and the system becomes an archive. Obviously the goal of ubiquitous assessment becomes moot [Nielsen 90]. Thus, one last critical requirement of an appropriate system is attention to the difficult issue of seamless interoperability. In order to understand how this is most likely to be achieved, it is useful to consider that the tradeoffs between the two approaches to creating ILS products are similar to those found in the software industry between the Integrated Development Environments (IDE) and Best-of-Breed (BoB). IDE means that a whole client/server application will be developed using a single development tool, including the front-end and middleware components. The BoB approach means that separate tools from different vendors are used for functions such as the front-end screens, middleware, and database access [Dolgicer 96].

A successful example of where the BoB approach has been used recently in the computer industry is Netscape's support of NetDynamic's NetObjects Fusion and Symatec's Visual Cafe in addition to Netscape's own LiveWire for building WWW sites. Netscape chose to resell these tools in boxes endorsed by their brand name, they provided front-line support for their products, and they worked with vendors to make sure the products worked together well. This example suggests that products should be compared on their vendor's interest in documenting and sharing standards with third party module developers, and also working with other companies to ensure seamless interoperability of standard modules.

The examination of a WWW-based learning system's architecture and support policies is a significantly different approach than using extensive "feature checklists" for comparing products to determine their strengths and weaknesses. Products that appear to have superior assessment capabilities based on "checklists of features" can have inferior architectures for supporting a wide range of assessment capabilities over time, while products with the most robust architectures and policies for supporting "open learning systems" can have fewer built-in features. If a product with an adequate architecture and support policy only has one or two high quality demonstration modules, it could easily be judged as "lacking many features" on a comprehensive "features checklist." This demonstrates why traditional approaches for comparing WWW-based learning systems with "feature checklists" could lead to choosing less effective packages. Individual features should be able to come and go, while the infrastructure remains the same.

6 An Example Comparison: Toolbook II and Authorware

The remaining task is to demonstrate how to determine the degree to which current WWW-based learning systems support a broad variety of embedded assessment features. For the purpose of this article, the discussion will be limited to a brief comparison of Asymetrix Toolbook II and Macromedia Authorware, two of the most popular commercial products. Rather than treat the following statements as definitive conclusions about the differences between specific products and companies, consider them illustrations of how differences appear when the criteria include the architectures and policies critical to long-term support for a wide variety of embedded assessment features.

The first point to make about Toolbook II and Authorware is that they both can be used in configurations with the technical characteristics necessary for accommodating a variety of assessment capabilities. Specifically, they can be used in a system that accommodates a core of reliable integrated management tools, one or more modules with instruction and assessment, standard database connectivity, and an acceptable level of attention to permissions and security. The difference between these products is how the technical architecture can be achieved.

Asymetrix ToolbookII is a family of commercial products that together provide a modular architecture similar to that described in this paper. Assistant and Instructor are desktop development environments for creating instructional modules and corresponding assessments. Librarian is the component in the suite that is used to manage and administer a course on the server (CMI Core Module). Asymetrix is concerned with interoperability, but this focus may be a function of the need to make the company's own expanding suite of products work together.

For example, Asymetrix recently bought AimTech's "IconAuthor" to expand the capabilities of it's ToolBook II family of products. IconAuthor includes support for generic computing standards such as Dynamic Data Exchange (DDE), Dynamic Link Libraries (DLL), and open API. While this might allow some products from other vendor's to be incorporated into the Toolbook II architecture, this is not very dependable since it is an artifact rather than a deliberate strategy to achieve the objective of interoperability. No reference was found to show that Asymetrix has actively cooperated with other companies or standards groups like the AICC to achieve interoperability between their own products and a variety of other WWW-based learning systems.

Macromedia's main product in the WWW-based learning systems market is "Authorware," a long time leader in the space of "classical" authoring software. In contrast to Asymetrix, the following points illustrate that Macromedia is actively building an architecture and implementing policies supportive of an "open learning system" that can support products from other companies:

• Macromedia bought the company Solis and its product "Pathway," which is one of the leading products designed to function as a "CMI core module" in an "open learning system" architecture. In one press release, Macromedia claimed that it would use the renamed product "Pathware" to provide some compatibility with Asymetrix Toolbook and Microsoft PowerPoint applications [Macromedia 97].

• Macromedia and WBT Systems have agreed to "work closely to offer a high level of integration" between WBT Systems' TopClass Server and Macromedia's authoring tools through the use of the AICC interoperability standard [Duff 97]. This particular alliance is even more significant when one considers that TopClass is often considered as a separate major product on its own in many reviews of major WWW-based course delivery systems.

This sample comparison between two leading commercial WWW-based learning systems shows that Toolbook II requires a tightly interwoven set of products from

Asymetrix, while Authorware is focusing on connecting products from a number of vendors. Thus, Macromedia currently appears to be the company most likely to provide a product that will accommodate a wide variety of instruction and assessment approaches over time through the incorporation of interoperable modules from other companies. On the other hand, Asymetrix appears to be more likely to remain an "insular" set of products from a single company. Use caution in interpreting the statements over time, because the market for WWW-based learning systems is expanding and volatile. Factors such as financial earnings and changes in company leadership can have an effect on these issues. If the objective is to choose a product, look for up-to-date descriptions of the products and watch press releases for evidence of activities such as architecture changes or strategic partnerships.

While there can be no certainty for now, either Toolbook II or Authorware could become the most widely adopted WWW-based learning system over the next few years. On the other hand, there are other products that could eventually function as a defacto standard across many educational institutions. In the face of such uncertainty, it is important to encourage as many companies as possible to accommodate architectures and policies conducive to supporting the full potential for improvements in assessment with WWW-based learning systems.

7 Conclusion

The most important role the computer may play in education could be contributing to the ubiquitous use of assessment for the improvement of instruction, rather than just serving as a better tool for delivering instruction and assessment to learners. In order to realize this potential, the newly emerging WWW-based learning systems need to support a very wide range of embedded assessment features. They should be able to include a variety of modules with tightly coupled and congruent instruction and assessment components for both classical and innovative approaches to instruction. The systems should also support assessments of a range of instructional variables for a variety of audiences for use in a cycle of continuous improvement. In order to support these demanding assessment capabilities, the architectures of emerging WWW-based learning systems must include a core of reliable integrated management tools, one or more modules with instruction and assessment, standard database connectivity, and an acceptable level of attention to permissions and security.

It is clear that no one company will be able to adequately address all of the possibilities for assessment in WWW-based learning systems, so it is critical that WWW based learning systems have "open system" architectures and company policies for cooperating with other companies to support interoperable modules. It is particularly important to note that products which might appear to have superior assessment capabilities based on "checklists of features" can have inferior architectures for supporting a wide range of capabilities, while products with the most robust architectures and policies for supporting "open learning systems" can appear to have fewer features. If the products that become the most widely accepted do not have the features, architectures and policies needed for the widest possible instruction and assessment capabilities, the computer learning environments of tomorrow will look like the traditional educational environments of yesterday, and potential opportunities for innovation will be lost.

References

[AICC 97] "AICC: Guidelines and Recommendations, AGR 007 - Courseware Interchange."; Aviation Industry CBT Committee (AICC), CMI Subcommittee (1997).

[Allan, Starry and Wright 73] Allan, R., Starry, J.O., and Wright, G.L: "An automated instructor and course appraisal system"; Educational Technology, 13 (May)(1973), 61-64.

[Anderson and Draper 91] Anderson, A. and Draper, S.: "An Introduction to Measuring and Understanding the Learning Process"; Computers Education, Vol.17, NO. 1, Pergamon Press, UK (1991), 1-11.

[Assessment Systems Corporation 97] Assessment Systems Corporation: "Testing Software and Measurement Books"; Assessment Systems Corporation, St. Paul, MN (1997), 7.

[Baker, Herman and Gearhart 96] Baker, E.L., Herman, J.L., and Gearhart, M.: "Does Technology Work in Schools? Why Evaluation Cannot Tell the Full Story"; In Education and Technology: Reflections on Computing in Classrooms, C. Fisher, D.Dwyer, and Yocam, K. (Ed.s) Apple Press, Jossey-Bass Publishers, San Francisco, CA (This is the ACOT Reflections) (1996), 185-202.

[Bennett and Hawkins 94] Bennett, D. and Hawkins, J.: "Alternative Assessment and Technology"; Educational Media nad Technology Yearbook, D. Ely, and B. Minor (Eds) Bold 20, Published in Cooperation with ERIC Clearinghouse on Information and Technolgoy and the Association for Educational Communications and Technology (1994), 151-154.

[Bernstein 88] Bernstein, M.: "The bookmark and the compass: Orientation tools for hypertext users"; ACM SIGOIS Bulletin 9(4), October, (1988), 34-35.

[Clark 94] Clark, R.E.: "Assessment of Distance Learning Technology"; In Technology assessment in education and training, E.Baker and H.O'Neil, Jr. (Eds.) (1994), 64.

[Derry, Seibert, Starry, VanHorn and Wright 74] Derry, J.O., Seibert, W.F., Starry, A.R., VanHorn, J.W. and Wright, G.L.: "The Cafeteria System: A New Approach to Course and Instructor Evaluation"; Measurement and Research Center, Purdue University, IN (IRB 74-1) (1974).

[Dolgicer 96] Dolgicer, M.: "The Meaning Behind the Best-of-Breed Approach"; IT/IS BackOffice, February Issue, Microsoft Press.(1996), 21.

[Duff 97] Duff, C.: "WBT Systems Teams with Macromedia to Broaden Support for Open Standards and to Develop a High Level Integration between their Respective Web-Based Technologies"; WBT Press Release, WBT Systems, San Francisco, CA (1997).

[Dynes, S., Litchfield, L., Curtis, K. and Chiquito, A. 97] Dynes, S., Litchfield, L., Curtis, K. and Chiquito, A.: "The Case for Database Support in Multi/Hypermedia Application Development and Use"; Proc. Ed-MEDIA'97, AACE Publishing, Calgary, Canada (1997).

[EvNet 96] EvNet: "EvNet Letter of Intent; Network for the Evaluation of Educational Technology Strategic Research Networks in Education and Training, Social Sciences and Humanities"; Research Council of Canada (1996).

[Feurzeig 87] Feurzeig, W.: "Algebra Slaves and Agents in a Logo-Based Mathematics Curriculum"; In Lawler, R. W. & Yazdani, M. (Eds.). Artificial Intelligence and Education, (Vol. 1). Norwood, NJ: Ablex (1987).

[Flagg 90] Flagg, B.N.: "Formative evaluation for educational technologies"; Hillsdale, NJ: Lawrence Erlbaum (1990), 235-236.

[Hammond and Allinson 89] Hammond, N. and Allinson, L.: "Extending Hypertext for Learning: An Investigation of access and guidance tools"; In People and Computers V A. Sutcliffe, and L. Macaulay (Eds.) Cambridge University Press, Cambridge (1989).

[Harel and Papert 90] Harel, I. and Papert, S.: "Software Design as a Learning Environment"; Interactive Learning Environments 1990), 11-32.

[Hawkes 97] Hawkes, M.: "Evaluating School-Based Distance Education Programs: Some Thoughts About Methods"; Evaluation and Policy Information Center, North Central Regional Laboratory (1997).

[Henderson 91] Henderson, J.: "Designing Realities: Interactive Media, Virtual Realities, and Cyberspace"; Multimedia Review, 1(2) (1991), 47-51.

[Hiltz, S. 90] Hiltz, S.: Evaluating the Virtual Classroom; In Online Educaton: Perspectives on a New Environment, L.Harasim (Ed.), Praeger, New York.(1990), 133-183.

[Hopper and Lawler 91] Hopper, M.E. and Lawler, R.W.: "Pre-Readers' Word Worlds: Results of Experiences with Young Children and New Directions"; Poster presented at 1991 International Conference on the Learning Sciences (Thirteenth Annual Meeting of Cognitive Science Society), Northwestern University, Evanston, IL, (August) (1991).

[Hopper 93] Hopper, M.E.: "Educational Courseware Production in Advanced Computing Environments" [Abstract]; In H. Maurer (Ed.), Proc. AACE Ed-MEDIA'93 - World Conference on Educational Multimedia and Hypermedia (1993), 607.

[Knussen, Tanner and Kibby 91] Knussen, C., Tanner, G. and Kibby, M.: "An Approach to the Evaluation of Hypermedia"; Computers Education, Vol.17, NO. 1, Pergamon Press, UK (1991), 13-24.

[Landow 89] Landow, G. P.: "Course Assignments Using Hypertext: The Example of Intermedia"; Journal of Research on Computing in Education (1989), 349-365.

[Lawler 82] Lawler, R.W.: Designing Computer Based Microworlds; Byte Magazine (August) (1982).

[LeBold, Montgomery and Ward 90] LeBold, W.K., Montgomery, R.E., and Ward, S.K.: "Computer and Self Concept Development of U.S. Engineering Students"; ASEE-IEEE Frontiers in Education, Vienna, Austria (1990).

[LeBold, Budny and Ward 97] LeBold, W.K., Budny, D. and Ward, S.: "Using Self Concepts in Placement and Evaluation"; ASEE Annual Conference, Milwaukee, WI (1997).

[Levin, Kim and Reil 90] Levin, J.,Kim, H. and Reil, M.: "Analyzing Instruction Interactions on Electronic Message Networks"; In Online Educaton: Perspectives on a New Environment, L.Harasim (Ed.), Praeger, New York (1990), 185-213.

[Macromedia 97] Macromedia Inc.: "Macromedia Announces Aquisition of Solis, Inc. Leader of Open Sysems for On-Line Learning Macromedia Pathware offers open solutions for on-line learning"; Macromedia Press Release, Macromedia Inc. (1997).

[Maddux and Willis 93] Maddux, C. and Willis, J.: "Integrated Learning Systems: What Decision-Makers Need to Know"; Ed-Tech Review Spring/Summer (1993), 3-11.

[McCollum 97] McCollum, K.: "A New Industry Sprouts Up to Help Professors Put Courses On Line"; Chronicle of Higher Education, October (1997), A33-A34.

[Newby, Stepich, Lehman and Russell 96] Newby, T., Stepich, D., Lehman, J. and Russell, J.: "Instructional Technology for Teaching and Learning: Designing Instruction, Integrating Computers, and Using Media"; Printice-Hall, Inc., Englewood Cliffs, NJ (1996), 301.

[Nielsen 90] Nielsen, J.: "Evaluating Hypertext Usability"; In Designing Hypermedia for Learning, D. Jonassen and H. Mande (Eds), Springer, Heidelberg(1990), 147-168.

[Oblinger 95] Oblinger, D.: "Transforming the Academy to Improve Delivery of Services: Redesign for Reallocation"; IBM White Paper, IBM Corporation (1995).

[Odell 97] Odell, J.: "WBT Systems Broadens Presence in Publishing Market: The McGraw-Hill Companies' Educational and Professional Publishing Group Licenses WBT Systems' Topclass Technology for the McGraw-Hill Learning Architecture (MHLA)"; WBT Press Release, WBT Systems, San Francisco, CA (1997).

[Papert 80] Papert, S.: "Mindstorms: Children, computers and powerful ideas"; New York, Basic Books (1980).

[Reeves 92] Reeves T.: "Effective Dimensions of Interactive Learning Systems"; Invited Keynote Paper Presented a the Information Technology for Training and Education Conference (ITTE'92), Queensland, AU (1992)

[Scriven 94] Scriven, M.: "Evaluation as a Discipline"; Studies in Educational Evaluation, Vol. 20 (1994), 147-166.

[Stager and Mueller 91] Stager, S. and Mueller, D.: "Computer Use in Classroom Testing"; In Educational Testing: Issues and Applications, Kathy E. Green, (Ed.), Garland Publishing, Inc. NY (1991), 245-277.

[Stufflebeam 73] Stufflebeam, D.: "Toward a Science of Educational Evaluation"; The Educational Technology Review Series, Number Eleven: Evaluation of Education, Englewood Cliffs, NJ (1973), 20-27.

[Thorndike 71] Thorndike, R.: "Educational Measurement for the Seventies"; In Educational Measurement, 2nd Ed., R. Thorndike (Ed.), American Council on Education, Washington, D.C. (1971), 3-15.

[WorldWired 96] WorldWired, Inc.: "Plato Reborn!"; Training with Multimedia, Vol 2, No6 (1996).

[Yankelovich, Meyerowitz, and VanDam 85] Yankelovich, N, Meyerowitz, N. & VanDam, A.: "Reading and Writing the Electronic Book"; IEEE Computer, 18(10) (1985), 15-30.