
Assessing the Virtual Classrooms: A Progress Report

Summer, 1995: Volume V, Number 2

Robert W. Tucker

In 1991, I devoted this editorial space to the subject of assessment in the distance learning environment [Virtual Assessment, AAF, I(3)]. I expected that innovation was only a matter of months away. Four years later, the only way virtual has worked its way into the discussion of assessment is to note that virtually no progress has been made. Given the need, opportunity, importance and cost factors, there are few explanations for this serious lack of progress. [Note: The concluding article in the series on assessing cognitive achievement is awaiting reviews of new statistical software.] - Ed.

Computer-mediated distance learning (CMDL) is an uncommonly bright star on the horizon of higher education innovation. According to a recent study by CCA and Associates, 30 percent of colleges and universities have at least one distance learning program in place and an additional 28 percent are planning such programs. According to Investor's Business Daily, more than 70 colleges and universities are offering online postsecondary education.

The University of California at Los Angeles the nation's largest college-based provider of continuing education, is pilot-testing its approach to CMDL. I expect to see the roll-out of a large-scale program from UCLA by 1996. In 1989 the University of Phoenix began a CMDL program with fewer than 100 students. By spring 1995 enrollment had grown to more than 1,100, mostly business students spanning the globe. Programs in Canada, such as those developed by Ontario Institute for Studies in Higher Education, Great Britain and Australia are also growing rapidly, while many on-ground programs show steady or declining enrollment. By the year 2000, the CMDL market, once viewed as a poor stepsister to on-ground education, will be crowded with newcomers and will have become a mainstream enterprise.

Benefits of Computer-Mediated Distance Learning

As with any social innovation that challenges a received view, there is contention and some confusion as to the value of CMDL programs. Being hooked to the rocketship of technology has been a mixed blessing. It has ensured rapid progress that outpaces mature understanding. Most of the construct development work surrounding CMDL has not evolved to the level of central principles and many empirical claims are based on scant, potentially idiosyncratic, evidence. Nonetheless, there are a few clear benefits to be found. Some of these benefits apply to all CMDL, but most apply more to the virtual time CMDL environment. I will suggest one more

benefit that holds very important implications for assessment and academic quality management in the CMDL environment.

Access. Perhaps the most agreed-upon benefit in the CMDL environment is that of access. CMDL programs can be delivered to students who live in remote and inaccessible places and for whom there are no practical educational alternatives. They can be delivered to students whose physical movements are impaired or otherwise restricted. They can be delivered to captives of the urban jungle or people whose jobs take them out of town most of each week. They can even be delivered to students who simply prefer to learn from the home, office or hotel room.

Access is especially great in the so-called asynchronous environment (the term virtual time is suggested as more appropriate) in which students are freed from the constraints of place and time when they interact across a period of one or more real days. Innovation in curriculum for CMDL and in store-and-forward computer conferencing software have contributed to the solidification of this benefit. Nonetheless, the benefits of access have yet to be thoroughly explored, especially the empirical claims associated to it.

Equity. Designers, instructors and students in the CMDL environments report that the virtual classroom is a more level learning field than the physical classroom. The underlying premise of these various claims may be that the virtual classroom removes some dimensions of classroom interaction which are commonly judged to be secondary, superficial or irrelevant to the core business of learning.

Dimensions such as physical appearance gender, body language, showmanship, shyness and competition for limited discussion time are minimized or eliminated in the virtual classroom. Many of these factors remain in the real time computer-mediated classroom; real time computer environments typically heighten competition for discussion space and place a premium on keyboard skills. Dimensions such as thoughtfulness, staying on point, precision of word choice, creativity and the ability to build upon preceding ideas may be enhanced in virtual time CMDL. One general factor seems clear from current evidence: shy, inhibited and deliberative learners are on a more-or-less equal plane of access in the CMDL classroom.

Potency. Most participants I have worked with find the virtual CMDL environment to be more compressed and potent than other delivery methods. Experienced instructors and students develop pithy and highly readable communication styles. Curriculum, lectures, assignments and learning group projects tend to be very well organized or they do not stand the unforgiving test of use. Slackers and the marginally knowledgeable (whether students or instructors) tend to be highly visible to the remainder of the learning group; both the quality and the quantity of contributions stand out.

That these and other educationally important differences exist should not be surprising, as vocal and keyboard communication represent significantly different cognitive activities. Different

portions of the brain are active, and the syntax and semantic content of speech and the written word are measurably different.

Cost-effectiveness. While technological advancements and potential educational benefits have provided a foundation for the growth of CMDL programs, economic factors may prove to be the ruling reason for their growth. (In this group I am specifically excluding the large public universities who, for strictly financial reasons, hope to leverage even further their irreparably corrupt educational models in which 350 students learn writing from a teaching assistant for whom English is a second language.)

Distance learning is increasingly seen as a means to increased educational productivity and cost control. The irony in this view is that only a few forms of distance learning actually hold potential for significant cost savings if one is to maintain the low student to teacher ratio (10:1 to 16:1) believed necessary to effective education in the CMDL environment. In all probability, only virtual time CMDL offers true economic advantages with small classrooms. Full cost allocation pricing for satellite education, for example, shows it to be more expensive than traditional on-ground education.

Scrutability. In many respects, an adult-centered education model for the physical classroom need not be altered to accommodate the virtual time CMDL classroom. Small learning groups can interact in the expected ways among and within themselves and with faculty. Syllabi, assignments, papers, documents, reports, graphics, sound, discussions and results of "offline" experiments and applications can move easily back and forth across the phone lines. Most solitary activity can be managed offline with store-and-forward conferencing technologies. In just these ways, the mass storage device of students' computers becomes a virtual map of the educational experience from the student's perspective. In retaining operational and curricular structures, as well as all transactions, the mass storage device of the university's central file server becomes a virtual map of the entire educational environment. This observation holds important implications for assessment.

The CMDL classroom is routinely a far more scrutable educational environment than the physical classroom. The virtual classroom's electronic data storage, retrieval and exchange system (i.e., the text of student and faculty transactions, communication logs, file structures and information presentation algorithms that exist on the file server's hard disk drives) represent concentrated, structured and highly accessible artifacts of the learning transactions. These artifacts can be retrieved, analyzed and reported via highly economical, automatic processes operating in the background of the communications system. In contrast, the human and material costs of capturing learners' transactions in the physical environment are so great that systematic studies are infrequent and limited in scope, and they require major funding and years to conduct.

The extremely low incremental cost of automated virtual classroom assessment makes it an attractive bargain to those who must conduct assessments on a low budget. Automated virtual assessment offers the potential for greater precision and reliability in assessment protocols; it is

also inherently less intrusive than most physical assessments and can adapt its procedures more or less on the fly, something very difficult for traditional assessment research. In fact, the potential of virtual assessment is so great that one thinks immediately of the need to carefully consider the issues of informed consent of any such assessment program.

CyberQ@: An Integrated Virtual Process and Outcome Assessment System

CyberQ is a software and hardware-based approach to adaptive assessment and academic quality management developed by InterEd for Apollo Group. The heart of this system is a conceptual model that can be adapted to most platforms, including DOS, Windows and UNIX. The automated assessment model outlined here is essentially the same as the model outlined in the Fall 1991 issue of AAF. While this particular system is proprietary, any institution with the resources and a commitment to assessment and continuous quality improvement can develop a similar system.

Functional Components

The CyberQ system measures an array of basic-to-complex or lower-order to higher-order educational constructs captured by the CMDL program's server and conferencing software. Participants (students and instructors) are aware of the system primarily through its automated adaptive administration of questions to assess academic structure, content, process and outcome. The administration of these intrusive questions can be random or they can be triggered by the achievement of specific academic milestones. For example, a "three minute assessment of achievement in the major field" can be triggered by the completion of a lesson plan or by the student's first login each week. Assessment of affective impact can be triggered by the students whenever they achieve an insight into the deeper, durable value of the material under study. Even while these higher-order measures are taking place, the system is gathering data at more basic levels. The system's key measures are outlined below.

Transaction Profiling Software. An array of easily obtained and profiled measures falls within the category of necessary but not sufficient measures to the assessment of academic quality. Among these are measures pertaining to the gross movement of electronic information. Ratios of incoming-to-outgoing bytes and profiles of faculty-to-student, student-to-faculty and student-to-learning-group ratios provide gross quantitative measures of the exchange of information in the CMDL classroom. If the CMDL environment creates learning groups, as it generally should, individual data can be aggregated to the group level where they are tracked and profiled as an indicator of the functioning of groups.

Of slightly greater intrinsic value than the above are derived measures such as query response times and rates. Whether or not an instructor responds to a student's query and how much time elapses between the query and the response is certainly an academic quality management and customer service issue. Information of this kind is organized along student-to-faculty, faculty-to-student, student-to-administration and faculty-to-administration relations.

It is important to view data of the above kind in the larger, longer term and to avoid the kind of reductionistic arguments that open this kind of analysis to legitimate criticism. One can, for example, conclude only a little about a student who forwarded 24,000 bytes of information in response to an essay assignment and forwarded this data when requested. However, additional conclusions may be warranted about a student who forwarded only 350 bytes in response to the same assignment and did so 10 days late; the same is true if one has a two-year profile on 100 instructors showing that the average response time between students' question and instructors' answers is 11.5 hours with a standard deviation of +2.5 hours. What action is warranted for the lone instructor with a 9.5 day average response time? Can one be certain that this instructor is cavalier with respect to his or her responsibility to students? Absolutely not. Is one justified in further investigation leading to a resolution? Absolutely.

Syntax Profiling Software.

Automated syntax analysis of writing can be accomplished through the computerized application of such simple means as Fleisch indices, in which syllable, word, and sentence length are counted. A more useful strategy has been the use of grammar and syntax checking engines that use a combination of lookup tables and expert system rules combined with advanced sentence parsing techniques to profile individual writing styles. While none of the software engines we have tested will do a reliable job of distinguishing among various levels of acceptable writing, most will do a fair job of dividing text into groups of unacceptable and acceptable writing structure. For the many students whose writing falls on the borderline between acceptable and unacceptable, this kind of syntax analysis will, over time and with sufficient samples, discern academic progress or the lack of it.

Michael Scriven has suggested that the routine use of grammar and syntax profiling software might be one way to respond to the problem of establishing students' identity in the CMDL classroom. Recently, Kathy McGuire, UCLA's project manager for distance learning, suggested to me that the problem of establishing student identity might be met by a hermeneutic approach in which a student's "write-print" could be constructed from one or two of the student's first pieces of writing in an online course. (The three of us are discussing this issue and invite your input. Join the discussion at the AAF e-mail address.)

Comment Profiling. The automated collection and profiling of CMDL transactions into primitive meaning relations offers great promise for assessment in the twentyfirst century distance learning environment.

My optimism for this approach is predicated on more than theory. Since 1989 I have been collecting, analyzing and profiling opened comments obtained from student end-of-course surveys. Based on the study of over one million comments, it is clear that these data possess superior attributes in three important areas when compared to data obtained from the scaled portions of the end-of-course surveys.

Comment data have greater salience for the respondent, their semantic content has greater fidelity to the respondent's true issues and comments represent issues of greater significance to the respondent.

The CMDL environment offers a significant advantage to comment profiling in that students' comments do not need to be translated from handwritten form on paper and pencil surveys. Comment profiles are of great value in assessing instruction, curriculum, learning environment and administrative support services. Contrary to one prevailing school of thought, comment profiles also show students to be responsible and constructive critics of their learning environment.

Predicate Analysis. On another level, whole text from the CMDL environment (assignments, reports, group dialog, instructor-to-student and student-to-instructor transactions) are subjected to what for lack of a better term, I am calling predicate analysis, in which text is massively grouped into meaning relations. The most straightforward way to conduct such assessments is to derive counts and frequencies for the use of descriptive, applicative, evaluative and analytic terms. These tables are examined according to differences among course levels or curricular domains. Data tables from accounting classes will have terms and frequencies of use that are immediately discernable from the data tables of classes on Renaissance literature.

A less reliable but potentially more valuable subset of this kind of analysis is the grouping of terms by associated value terms. This approach produces counts and frequencies of technical terms along with counts and frequencies of their pairing with negative and positive value terms. Relations among technical terms and value term valences are then mapped. In an economics class, for example, technical terms related to supply-side economics may be more often associated with negative value terms than a competing theoretical model.

By way of pure speculation, predicate analyses might be performed using an expert system with an empirically or normatively derived rule base. Even more complex, but still feasible, is the analysis of levels and kinds of group functioning. This could be done by embodying group process models into path analysis matrices. Neural networks are suited to this latter task but linear strategies will probably do as well, although they may require more experimentation. Less complex approaches, such as correspondence analysis, may also reveal interesting relationships.

Many of my colleagues feel that predicate analysis is unlikely to lead anywhere useful. Perhaps so, but when the potential benefits, however remote, are weighed against the costs, it would seem important to analyze the massive amounts of text being gathered in the CMDL environment. In time, these analyses will prove fruitful to understanding not only the distance learning environment but all educational transactions. Educators and pedagogists stand to gain new insight into cognitive stage and sequence, and we may yet derive learning taxonomies that are more sound than those currently guiding the production of curriculum. (The empirically derived structure of Bloom's taxonomy, for example, is substantially different from the formal structure.)

We are currently exploring the value of acquaintance categorization, a potentially promising new approach to text analysis described by Marc Damashek of the U.S. Department of Defense. This purely statistical method categorizes text according to sequences of consecutive characters. The advantages to acquaintance categorization are that it is language-independent, relatively easy to program and handles large amounts of text. Unlike predicate analysis, acquaintance analysis does not depend on meaning to efficiently group text by context and topic. To my knowledge, no one has yet applied this strategy to the CMDL environment.

Adaptive Assessment. All of the above measures take place without the explicit awareness of the participants (of course, participants should be made fully aware that the system is operating in tandem with their conferencing software and hardware). Another group of measures is taken by asking students and faculty to respond to specific questions throughout the progression of the CMDL course.

Some questions are asked at the time of login, while other questions are triggered by the passage of specific milestones in curriculum or measured learning outcome. In most cases, students and faculty have the option of skipping questions, with the understanding that doing so will increase the assessment burden in subsequent exchanges. This flexibility allows students and faculty to better adapt their educational activities to the other activities of the day. Some questions that assess achievement have time limits designed to assess unaided recall of subject matter knowledge, while others permit time to consult texts and reference materials.

The computer administered assessments include measures of: (a) achievement of course learning objectives, (b) student educational goals and goal fulfillment, (c) student affective disposition and growth, (d) student satisfaction with instruction, curriculum and the learning environment, and (e) faculty assessment of text, curriculum, general preparation of students and academic support services.

In most cases, adaptive assessment significantly reduces the length of the total question set. In all cases, it make the computer administered questions appear to be more intelligent.

Student demographics and marketing data are also gathered via this system. In most cases, question algorithms are adaptive (i.e., questions are adjusted to reflect answers to previous questions). The next question administered to a student who picks an incorrect answer to an achievement question is easier than the previous question. Students who indicate that they are single will not see a question asking for their spouse's name. In most cases, adaptive assessment significantly reduces the length of the total question set. In all cases, it make the computer administered questions appear to be more intelligent. Hierarchical Measures. While many of the above measures are narrow, each measure is a necessary condition or building block for some higher level of analysis. At this point it is not possible to predict which of them will form the most interesting associations with other measures. One can only experiment; all of this information must be creatively applied to explore, explain and refine the virtual learning environment.

Benefits of Automated CMDL Assessment

I must defer a more detailed discussion of the benefits of automated assessment systems to another issue of AAF. However, the following short list of the benefits of such systems should start the discussion.

The CyberQ system:

- ◆ provides a software implementation of a multi-trait, multi-method assessment strategy;
- ◆ provides virtually simultaneous assessment of multiple students at multiple levels;
- ◆ gathers data in ways that are largely transparent to students, faculty and staff;
- ◆ imposes minimal a priori theoretical assumptions on data acquisition processes;
- ◆ aggregates very large amounts of data for virtually unlimited secondary and meta-analyses;
- ◆ provides large amounts of data at very low cost per unit;
- ◆ tracks data indefinitely for longitudinal (pre/post) assessments;
- ◆ provides cost-effective and rapid administrative reports on academic mission-critical exception data;
- ◆ provides faculty and students access to data of value to them; and
- ◆ conducts background test-item analyses.

I would like to hear from those of you who have CMDL program(s) in place or who plan to develop such programs in the near future.

Robert W. Tucker

is President of InterEd

and the Phoenix Institute.