Preface

Section I: Higher Education

Chapter I. “Continuous Assessment for Improved Student Outcomes: Examples from WebCT’s Exemplary Course Project”

In the late 1990s a widespread instructional technology explosion occurred with the development and advancement of Internet and computer technologies. One scholar has referred to this period of time as being as important and significant to learning as was the invention of the Gutenberg press, turning the educational world upside down in so many ways. This phenomenon was brought to the common educator by instructor-friendly learning content management systems (LCMS) such as WebCT, Blackboard, eCollege, and many others. This chapter reveals the remarkable impact that online assessments and measurements embedded in LCMS tools are having on teaching and learning and leaves the reader with a desire for much more than the word limit on chapters in this volume allows. It is clear from this chapter and the many examples identified from the WebCT Exemplary Course Project that the innovative and best online assessment strategies made possible by LCMS in recent years are redefining not only the way instructors assess but also, more significantly, the way teachers teach and students learn. These authors briefly introduce many of the numerous online assessment features available in WebCT and illustrate the best practices through examples provided by instructors now using WebCT for online assessment. One of the most notable and unique features of an online assessment system such as WebCT is its capacity to provide immediate and quality feedback to the learner on quizzes and exams and also allow him or her to self-check exercises introduced throughout the LCMS environment. The authors acknowledge that the dream of assessment being more closely integrated with learning and even being indistinguishable from each other is a step closer because of LCMSs such as WebCT.

Chapter II. “Coordinating a Distance Learning Software Rollout with the Needs and Visions of a Mature Test Organization: Political and Technical Lessons Learned”

Any university or college seriously considering the adoption of a large LCMS (e.g., WebCT or Blackboard) with its online assessment and measurement capability for not only its campus and testing center but also its distance education programs will
benefit from this chapter. The sometimes difficult and costly lessons learned by one university promise to inform the strategic planning efforts of other like institutions, especially if they already have an established testing center. The competing needs and interests of faculty in the classroom, administrators in the testing center, and both in the distance education context are best addressed before, not after, the integration of an expensive, enterprise-wide LCMS for online assessment and measurement purposes.

Chapter III. “Assessing the Relationship between Learner Satisfaction and Faculty Participation in Online Course Discussions”

This chapter is written by senior administrators from one of the largest and most well known for profit international distance education institutions: Capella University. The brief history of this successful “dot edu” university-provided by the authors as context for their case study—is as fascinating as the case study itself. This specific case study focuses on the assessment and evaluation of the instructor and adult learner interaction in an online setting by using the Capella-developed Faculty Development Feedback System (FDFS). The authors explain how FDFS is used to assess online instructor immediacy within a course by measuring the frequency and quality of instructor and learner interaction in online course discussions. The authors also examine the intersection of FDFS results and course evaluation ratings with the surprising conclusion that the learner-content relationship may be more important than the learner-instructor relationship. This finding further validates a claim by distance education researcher Terry Anderson that “sufficient levels of deep and meaningful learning can be developed as long as one of the three forms of interaction (student-teacher; student-student; student-content) is at very high levels. The other two may be offered at minimal levels or even eliminated without degrading the educational experience” (Anderson, 2002, p. 4).

Chapter IV. “Authentic Online Assessment: Three Case Studies in Teacher Education”

This chapter reminds readers of the importance of aligning objectives, activities, and assessment. Three diverse teacher education programs at the Hong Kong Institute of Education undertook an initiative to update their assessment strategy by using technology-mediated tools unavailable to the teacher and student just a few years ago. This effort not only realized new insights about and experience with online assessment but also resulted in a redesign of the whole curricular package to ensure full alignment and integrity among all the pieces. The authors extol the benefits of online assessments, especially in promoting more timely and helpful peer collaboration and review. They note the value of using online assessments to better prepare students for high tech jobs, where online assessment and measurement are commonplace and to succeed in lifelong learning opportunities that will increasingly rely on online assessment and measurement strategies. The three case studies also showcase, through the eyes of students, the impact that formative-not summative-online assessments and feedback had on their learning process.
Chapter V. “Electronic Tools for Online Assessments: An Illustrative Case Study from Teacher Education”

The chapter highlights the creative use of common software tools by teacher education professors in a university course at Eastern Michigan University to facilitate and automate elements of online assessment involving student papers and portfolios while modeling the same for their students-soon to be teachers themselves. The authors introduce the use of rubric-building and-customizing software, databases, and the track changes feature in word processors with the help of a number of illustrative figures. Even though the tools described are common and accessible to both instructors and university students alike, their use, particularly in concert with each other, appears less common. Some benefits and advantages for use of these tools identified by the authors include more valid measurements, efficient grading, and modeling of online assessment methods to future teachers. The authors also note the importance of formative online assessments in helping students improve the educational product (e.g., paper or portfolio) and the instructor in improving the overall teaching, assessment, and feedback process.

Chapter VI. “Online Assessment of Foreign Language Proficiency: Meeting Development, Design, and Delivery Challenges”

The author of this chapter has taken a very complex subject involving an elaborate model of computer-adaptive testing for language placement and proficiency and explained the process and issues confronted in converting the exam to a Web-based format in a very lucid and engaging manner. The author’s endnotes are not to be overlooked either, as she defines in a simple and straightforward way three terms or phrases that help explain semi-adaptive or hybridized computer-adaptive tests developed specifically for this language testing context. This case study focuses on the adaptation of two less commonly taught languages (LCTL), Arabic and Russian, from a traditional to online format while introducing the reader to the many issues and challenges confronted by the team-some of them unique to online language assessment and measurement. As part of this project, the Center for Applied Linguistics and Georgetown University conducted an analysis of online assessment vendors by using predefined characteristics, for example, security, adaptability, price, and so forth, that will be helpful to other institutions looking for the right vendor fit. This chapter concludes with five pragmatic lessons that will generalize to other online assessment initiatives not exclusive to LCTL online assessment.

Chapter VII. “Using Web-Enabled Technology in a Performance-Based Accreditation Environment”

This chapter, written by five engineering faculty members from Penn State University, touts the role that two different online assessment systems played in helping their college prepare for professional engineering accreditation. Much of this volume, and most online assessment and measurement efforts themselves, focuses on the student and not the instructor; this case study features the instructor’s role and participation in online assessment and measurement systems. The first system developed by these engineers is clearly an example of using technology as a lever for
change to build consensus and foster democratization so that their accreditation effort is not perceived as only a top-down mandate. Because outcomes are the focus of new engineering accreditation standards, the 17 engineering programs of the 12 campus university system turned to a simple Web-based tool to build outcome consensus from a dispersed faculty. This online self-assessment instrument captured on a weekly basis three types of class level data from the faculty: learning goal(s), learning activities to support each goal, and performance summary. The authors speak of the benefits that accrued from this exercise, including the use of the aggregated data to make and build program level decisions and outcomes. They also remind the readers throughout the chapter about the importance of online assessment systems being scalable-stretchable and shrinkable-according to need. The second online assessment system developed focused on three data sources or measurements that would provide further evidence to the accreditation agency that outcomes were being met. Those three data targets included student performance on each outcome (as determined by the faculty member), faculty perception of course effectiveness, and students’ perception of their own degree of mastery at the outcome level. This triangulation of data clearly provides the multilevel, outcomes-based measurements critical for not just accreditation but, even more importantly, continuous improvement and advancement of student learning.

Chapter VIII. “Moving Beyond Objective Testing in Online Assessment”
The Scottish Centre for Research into online Learning and Assessment is taking online assessment where it has never gone before. Many of the questions that educational institutions are asking today about online assessment are in the past for this university research center, as they look into the future and beyond the constraints of today’s assessment theory and technology. Although the six authors of this chapter are constrained by space limitations, they still succeed in engaging those educators who read this chapter with an eye to the future of online assessment and measurement. This chapter emphasizes the importance of aligning multimedia teaching methods with multimedia assessment methods; it also challenges educators to not so quickly dismiss the capacity of online assessments to test higher order skills and highly symbolic subject areas (e.g., mathematics). This case study also provides some interesting background and history on where United Kingdom computer-aided assessments (CAA) have been and where the authors predict they are going. These authors and the research coming from their institute portend a day when rich multidimensional and authentic simulations/animations/presentations are commonplace in both formative and summative assessments, and where reporting and feedback along the instructional way is both rich and immediate. The overriding theme of this chapter is the need to blur the line between learning and assessment because technology allows a total integration of the same and a redesigning of the educational paradigm.
Section II: Elementary and Secondary Education (K–12)


The five authors of this chapter provide any state office of education considering computer-adaptive testing (CAT) in reading and math a helpful case study that examines large-scale CAT administration with suitable accommodations for disabled students. This study and pilot over 3 years touches on a number of online CAT issues through the accessibility perspective. These five authors are as conversant with the legal and moral responsibility of accommodating their disabled students as they are with the complexities of item response theory—the foundation of CAT. These practitioners and researchers alike have arguably taken CAT where it has not been before as they modify and adapt typical CAT delivery to accommodate content domain coverage and the atypical learner. This pilot exposes, as have other case studies in this volume, the importance of not overlooking technical requirements, best testing practices, and political nuances for rolling out a new testing paradigm on a large scale while accommodating students with special needs on a small scale. The authors categorize lessons learned from this experience under six headings: usability, comparability, development of the item pool, test security, state/local infrastructure, and buy-in from test consumers. Finally, the authors raise both important and provocative questions about the intersection of online assessment with federal/state/local regulations, including “No Child Left Behind.”


Computer-adaptive testing (CAT) in a K-12 smaller school district setting is the focus of this chapter. This small district successfully introduced this “new” type of online assessment using a relatively “new” type of testing theory—item response theory—and at the same time documented critical issues and questions that other districts need to address before they make a similar adoption. Although the authors do their best to keep this chapter as nontechnical as possible and demonstrate that CAT can be more efficient and reliable than traditional testing methods, some background in CAT would be helpful. However, even those without a rich background in CAT will benefit from a discussion of issues emerging from the case study. Although many of the benefits and drawbacks of CAT are similar to those of the more linear and traditional forms of online assessment, many are unique to this format: different levels of examinee familiarity with computers, shorter exam lengths, few items that are really easy or hard, lack of item review capacity, and examinees missing about half of the items before the algorithm brings the test to an end. This case study also emphasizes the important role of professional development and in-service in helping teachers and administrators better interpret results at the individual, class, and district levels and then make just-in-time curricular adjustments for their students. Furthermore, the authors posit the use of CAT data to identify trends in student achievement.
that can then be associated with best practices that are worthy of dissemination to other teachers. Finally, the authors give reasons why this CAT should supply just one of the measures of student progress and that its best use in the school district setting is “to support student learning by providing timely information to teachers in the classroom, and school administrators”.

Chapter XI. “A Computer-Adaptive Mathematics Test Accommodates Third Grade Students with Special Needs in the Pacific Northwest”

This case study examines the use of a computer-adaptive mathematics test administered to 250 third grade students in the Pacific Northwest with limited reading and writing skills. Out of fairness to the students and as a matter of integrity in reporting test results, it was important that no students who understood math but struggled with reading and writing be misrepresented in ability or otherwise left behind by the online assessment. The researchers endeavored to incorporate recent findings and recommendations from APA’s Task Force on Psychological Testing on the Internet in ensuring that this online assessment was in compliance. The authors examine issues of reliability and validity for this online assessment that include appropriate item development as well as what they call “structural development.” With item development, they employ “read-aloud and simpler vocabulary strategies”; with structural development, they go to great lengths to ensure consistent access and presentation of 10,000 learning assessment objects across an inconsistent computing environment, for example, varied platforms, browsers, screen sizes, color depths, and so forth. This chapter briefly introduces the many software tools used to develop the exam with some rationale for their selection and place in the entire process. The authors appear to have made every reasonable effort to anticipate and manage the many challenges and variables inherent in accommodating not only disabled students but also digitally disadvantaged schools. While the authors herald the benefits of online assessment and measurement, especially for providing accommodations that are otherwise not as readily available, they also acknowledge the phenomenon that sometimes occurs in solving one problem and then introducing a new one.

Chapter XII. “Designing an Online Formative Assessment that Helps Prepare Students and Teachers for a Summative Assessment: A Case Study—A Two Year Pilot Pairing an Online Classroom Assessment with a Statewide High-Stakes Test”

This case study showcases the role of online assessment and measurement in better preparing elementary students and their teachers for a mandated statewide test. Even though the statewide test is paper-based, the Gates Foundation and the Washington State Education Department piloted an online initiative to assess students along the way so that teachers could better respond to areas of concern in a more timely fashion. This experience highlights some of the advantages of online assessment and measurement, particularly in a diagnostic role, while at the same time exposing some of the practical and technical challenges that are frequently overlooked in the transition to more automated and computer-based testing environments.
Chapter XIII. “Online Assessment in the K–12 Classroom: A Formative Assessment Model for Improving Student Performance on Standardized Tests”

The importance of formative assessment along the instructional path is at the heart of this chapter. Much of this book (and much of assessment itself) focuses on the final or summative test that determines placement, certification, grade, and/or level. However, the real opportunity of online assessment may be in uniting its technical capacity with the instructional theory of continuous feedback in providing formative feedback to students and instructors along the way. Even though these case studies have been conducted in the K–12 context, the use and application of formative assessment methods similar to these generalize too many other instructional instances.

The three specific examples used to demonstrate the multiple student and instructor effects of using an online formative assessment system include (a) the No Child Left Behind (NCLB) initiative with 223 elementary schools in Kentucky and Tennessee, (b) 35,300 students in grades three through eight in Tennessee preparing for the Tennessee Comprehensive Assessment Program (TCAP), and (c) several elementary schools in Alabama identified as “priority” schools with some other brief examples of similar efforts in Kentucky and elsewhere. The authors also look ahead at the increasing need for and interest in the use of online formative assessment in K–12 schools nationwide; they identify reasons for predicted increases in the use of online formative assessment. They also include limitations of the current model, for example, not enough constructed response item types, and so forth. Finally, the authors briefly address the age old debate of “teaching to the test” in the context of these formative assessment systems.

Section III: Corporate and Government Training

Chapter XIV. “Online Assessment Distribution Models for Testing Programs: Lessons Learned from Operational Experience”

This first part of this chapter provides a brief but fascinating review of the evolution of online assessment and measurement since the 1970s. Many readers will be surprised to learn of the origins and milestones of computer-based or online assessment chronicled in this chapter. The author, vice president of testing services for Pearson VUE, then proceeds to define and describe three major categories of online assessment delivery entitled (a) controlled, (b) global, and (c) ubiquitous. The controlled model is as its names implies—very controlled. Online assessments that are controlled are usually administered by organizations that consider the results sufficient to make high-stakes decisions, for example, employment, advancement, and so forth. The characteristics of a controlled model generally include restricted access (as contrasted to global outreach), tight security, careful authentication, complete standardization, and a proscribed testing environment. At the other end of the delivery model continuum is the ubiquitous model (also known as Web based), and somewhere in between is the global model. The author emphasizes, regardless of model selection, the need for test owner, devel-
oper, and deliverer to work carefully through the many issues involved in the successful distribution of online assessments. He also shares some creative but effective data mining strategies used to detect possible security breaches and promote tighter security. At the conclusion of the chapter the author looks into the future and predicts more sophistication and capacity in all dimensions of distribution that will better enable online assessment sponsors to satisfy their goals with fewer of the current constraints.

Chapter XV. “From Paper and Pencil to Computerized: A Dental Board’s Perspective”

A large testing company, Thomson Prometric shares its experience in helping the Dental Board shift its certification testing approach for dentists, hygienists, and dental assistants from paper to the Web (computer based over secured networks and Internet based) after a security breach occurred in 2001. The author reviews previous concerns with online assessment, focusing on general access to computers, unfamiliarity with computers, and computer anxiety, and then points out how few, if any, of these historical concerns still exist today. This case study reminds any company looking to transition its assessment model of the current and relevant issues executives should consider in their analysis of whether they transition from paper to the Web, to what degree, and how quickly. Some of the online assessment and measurement issues addressed through this case study and discussion include security, a more authentic testing experience, exam availability and distribution, automated scoring, additional respondent data, item and form functionality, administrative efficiencies, technical considerations, data storage, and costs. After the Dental Board case study considered all these issues, one exigency was emphasized as probably more important than them all: “selling the idea to the candidates.” Change does not always come easy, and even though it is evident that online assessment is the wave of the future, many test takers would still prefer the former way.

Chapter XVI. “Assessment Elements in Web-Based Training”

The online training and assessment company, Imedia.it, was contracted by the U.S. Army to develop simulated instruction and assessment. This chapter recounts the experience of the army and Imedia.it in transitioning classroom training and assessment to an online setting. The specific instructional and assessment need that the army brought to Imedia.it was how to more efficiently train counterintelligence military personnel in effective interviewing techniques and strategies. Historically, this training module required such intense trainer involvement with learners that one instructor could train only two learners at a time, creating a tremendous training backlog. The new model of training and assessment reduced the trainer/learner ratio from 1:2 to 1:10–15 by using computer-based instruction and assessment. This performance-based training and assessment approach appears successful from a number of perspectives, including administrative efficiency, interrater reliability, and learner mastery; it also introduces creative and diverse approaches to maintaining online assessment security—a nagging challenge for online assessment and measurement. The case study also highlights the use of customized learning based on the results of diagnostic or pre-assessment exercises to create customized or adaptive learning for each soldier. Finally, the author examines the future prospects of merging CAT with
computer-enabled individualized and adaptive learning.

Chapter XVII. “The Seven C’s of Comprehensive Online Assessment: Lessons Learned from 36 Million Classroom Assessments in the Cisco Networking Academy Program”

In a day of corporate nondisclosure and secrets, this chapter is a welcome and refreshing addition to the growing body of knowledge concerning online assessment and measurement from a large corporation. Arguably, Cisco Corporation and its Networking Academy Program have had more experience in online certification assessment with its 36 million assessment instances than any other institution. The seven C’s described herein are best practices as much as they are overarching principles for anyone involved in online assessment and measurement, including certification testing. Although some people may consider the seven C’s as common sense, invariably one or two of the C’s have been overlooked in many institutional assessment strategies and have become the Achilles heel in an otherwise successful experience. The seven C’s are claims, collaboration, complexity, contextualization, computation, communication, and coordination. In context, each of these terms becomes a story unto itself and represents the critical and somewhat discrete dimensions of online assessment that, when combined, enjoy a synergistic and successful interaction. The authors share glimpses of the genius of their proven but innovative model by briefly introducing their use of networking—literal and figurative—to link instructional objectives to learning activities to assessment and then back; to provide a worldwide collaboration in instrument design and development, including language and cultural translation for over 150 countries and nine languages; and to just-in-time prescripted computations that enable the assessment professional to customize and interpret the results according to unique institutional needs. The third C, complexity, while referring to the capacity of technologies to accommodate new and more complex item types, for example, simulations and other performance-based assessments, also represents the many complexities and exciting opportunities associated with online assessment and measurement.

Chapter XVIII. “Feasibility Studies on Transitioning Assessment Programs from Paper and Pencil to CBT Delivery.”

Yes or no—should the paper-based assessments be converted to online assessments? What are the factors to be considered? The costs? Although some problems will be solved or at least mitigated and some efficiency introduced, will new problems and inefficiencies result? The experiences of these authors in conducting feasibility studies for clients considering the transition from paper to electronic are profound; this chapter promises assistance and insight not readily available to committees tasked with the transition from paper to electronic assessment feasibility study. As profound as these findings may be, they are just as frank and honest. The authors’ experience reveals that the transition actually costs more—at least net cost is more after accounting for net gains and losses. One of the biggest challenges introduced is determining actual costs of the current paper-based program to be used as a benchmark for comparison and analysis. Five phases for a feasibility study are discussed: Phase 1: Strategic Planning with the Sponsor, Phase 2: Extended Data Gathering, Phase 3: Integration of Data Gathered and Analysis of Costs and Timeframes, Phase 4: Integration and Interpretation of Data to Identify Deal Breakers, and Phase 5: Development
and Presentation of Feasibility Report. Threading their way through all these phases are stakeholder support and buy-in, which the authors emphasize as the most critical aspect of any feasibility study. Finally, three exhibits are included that are ready-made checklists or questions to gather information about the existing examination program, expectations of the stakeholders, and vendor characteristics that best match up with institutional objectives.

One of the authors of this volume may have best summed up the excitement, optimism, and outlook associated with this book and its discussion about online assessment and measurement: “We are just beginning to have a small idea of the real extent of the possibilities for online assessment.” Unquestionably, all the authors, researchers, and practitioners herein wish that they had had the benefit of this volume’s collective experience and knowledge before beginning their experiments and adventures with online assessment and measurement. It is the hope of these editors that this book series will encourage and inform the practical and scholarly inquiry and application of online assessment and measurement for many years yet to come.

Scott Howell
Mary Hricko
December 2004