The Association of Test Publishers
Innovations In Testing
Conference 2005

Computer-Based Testing's Not Just for Kids Anymore:

Developing Innovative CBTs for Adult Basic Education

April L. Zenisky and Stephen G. Sireci

Center for Educational Assessment
University of Massachusetts Amherst

Poster presented at the 2005 conference of the Association of Test Publishers, Scottsdale, AZ.
© Center for Educational Assessment, UMass-Amherst

Abstract

Two of the most-cited potential advantages associated with computerized test administration are the adaptive testing possibilities and the use of novel items and response formats. These benefits help to explain why computer-based-testing is increasingly finding considerable application in assessments for adult basic education (ABE) in the United States, where skills in the examinee population range from pre-literacy and pre-numeracy through GED. The purpose of this poster is to describe how some of these innovations are being incorporated into computerized reading and math tests now being developed for ABE in Massachusetts. Of particular note is the proposed three-stage design with adaptive-by-item and linear stages, and also the use of novel item formats to provide examinees with opportunities to demonstrate reading and problem-solving skills in real-life contexts.

Background

- Adult Basic Education in Massachusetts
  - Provide educational services for 20,000+ adults in basic literacy, numeracy, and HS equivalency
    - Both ESOL and native speakers of English
- Programs funded by state and federal government must document learning gains
  - Currently, paper-based tests from a national test developer are used

© Center for Educational Assessment, UMass-Amherst

Goals for the New Tests

- Aligned with Curriculum Frameworks for Massachusetts ABE
- Administered via computer (and Internet) if possible
- As short as necessary for good measurement of individuals – computer-adaptive if possible
- Using novel item formats where appropriate to measure student skills
- Incorporating Universal Test Design principles for accessibility to all students

© Center for Educational Assessment, UMass-Amherst

To accomplish these goals:

- Proposed design is a multi-stage test, with the following three stages:
  - Stage One: A “Locator”
    - Short section to help identify where (in difficulty terms) to start administering items, CAT-style
  - Stage Two: Traditional adaptive-by-item CAT
    - Gain a high level of precision in ability estimates
  - Stage Three: Linear, fixed set of items
    - Chosen based on estimated proficiency at end of Stage 2; will ensure content coverage

© Center for Educational Assessment, UMass-Amherst

To accomplish these goals:

- For test delivery, the Center is working with the UMass Computer Science department
  - We are adapting an online intelligent tutoring system for large-scale testing
- Teachers developed curriculum and test standards
  - Trained in item writing and item reviewing

© Center for Educational Assessment, UMass-Amherst

Proposed Design

Stage 1

Background:

Students and/or teachers answer few, short items about their familiarity and skills in content area

Stage 2

Adaptive:

~20 items selected as appropriate difficulty for different students

Stage 3

Linear:

Sets of ~20 items classified by level and varying by difficulty presented to students

=

End of Test

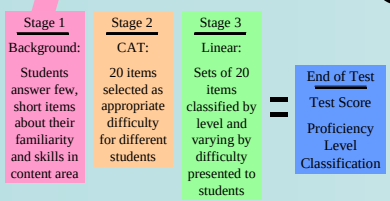
Test Score

Proficiency Level Classification

© Center for Educational Assessment, UMass-Amherst

Stage 1: Background

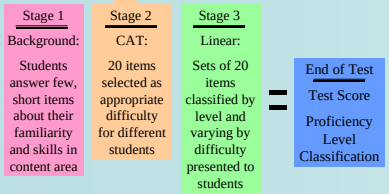
- Need to measure small gains over continuum of ability
  - Need precise measurement for individuals
- Stage 1 will help to find a starting place for the test
- Teachers may have a role
- Others?



© Center for Educational Assessment, UMass-Amherst

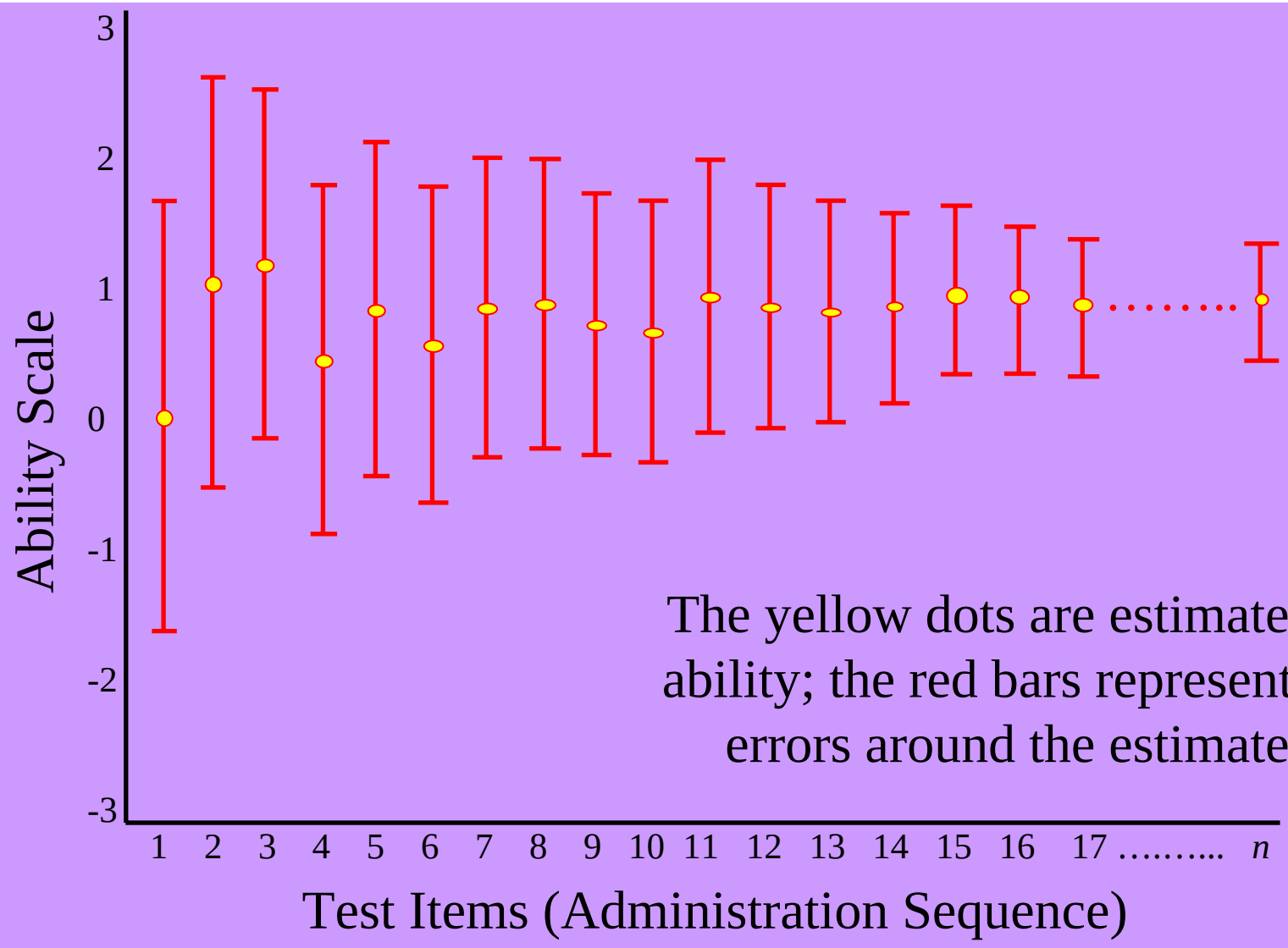
Stage 2: Adaptive

- In the second stage items are selected based on student responses
  - Each student is presented ~20 items chosen this way
- Less-able students will not be frustrated with many hard items (and vice versa)
- Test starts based on Stage 1 results



© Center for Educational Assessment, UMass-Amherst

Example: Estimation in Action



The yellow dots are estimates of ability; the red bars represent the errors around the estimates

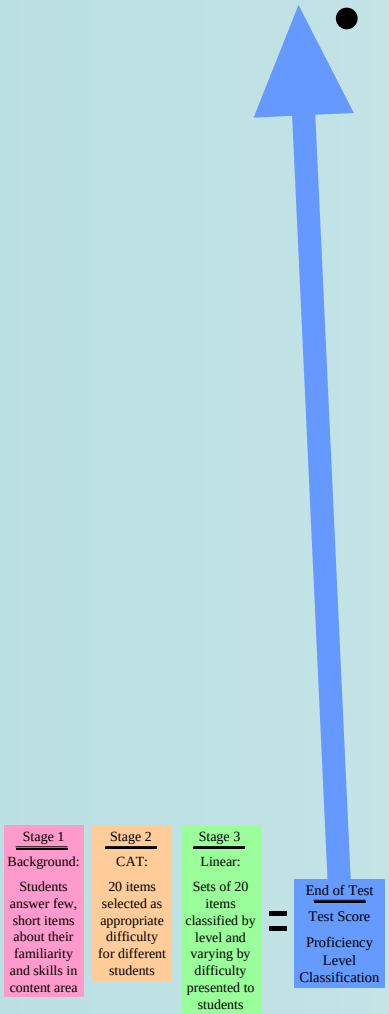
Stage 3: Linear

- In the third stage, students have been roughly assigned to levels
  - Here, each student receives ~20 items to ensure content coverage and allow for provisional Level classification decision

© Center for Educational Assessment, UMass-Amherst

End of Test

- When the test is complete, each student receives a scale score and a proficiency level classification linked back to the curriculum frameworks
  - Scale score range is 100-600
  - Whenever possible subscores to inform teaching and learning will be provided
    - Trade-off: unreliability vs. time on test



© Center for Educational Assessment, UMass-Amherst

Types of Items

- Multiple-choice
- Short Answer / Fill-in
- Innovative Formats
  - Can measure skills and abilities that are not appropriately measured by other approaches
- Regardless of format, most items are set in contexts adults may encounter in their everyday lives
  - Bus schedules, nutritional tables, etc.

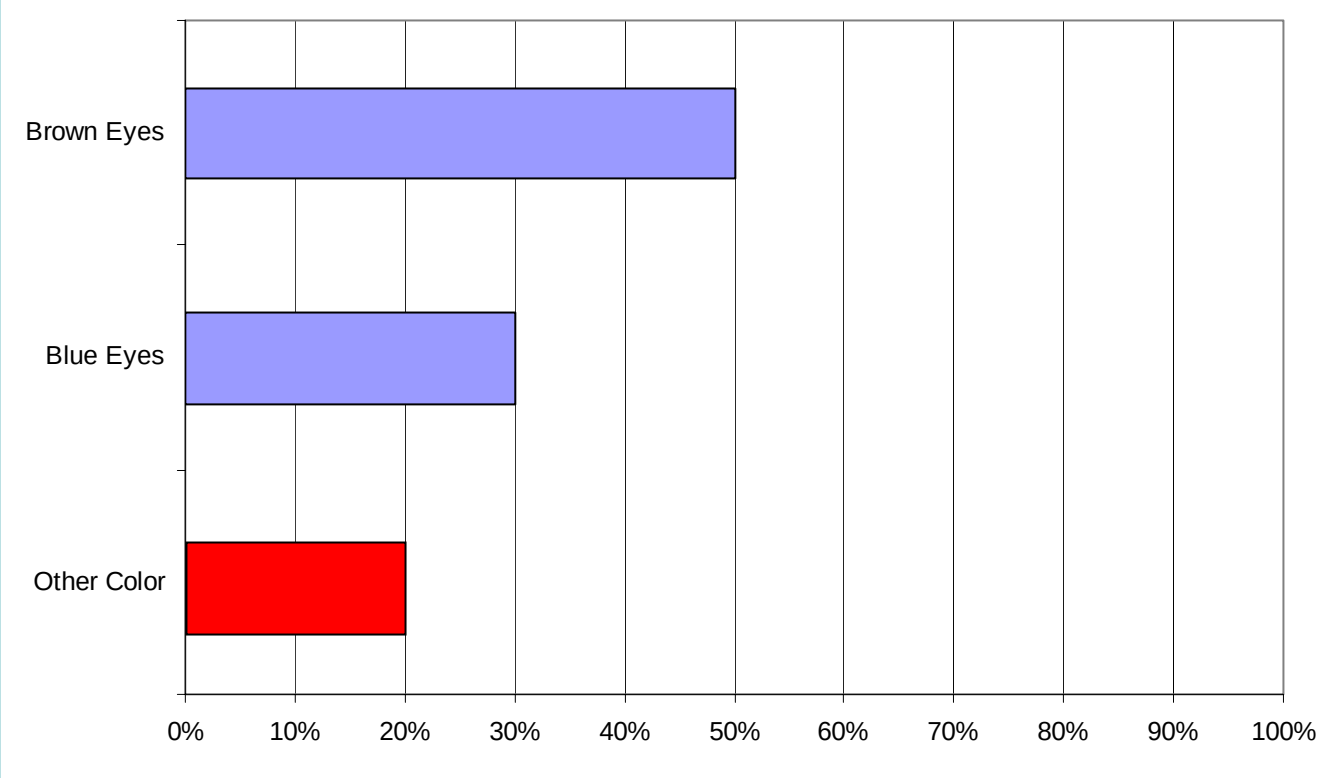
© Center for Educational Assessment, UMass-Amherst

Graphical Modeling

(Bennett, Morley, & Quardt, 2000)

Directions: Use the information in the table below to finish the bar chart. Drag the bar labeled 'Other Color' to show the percent of people in that category.

Eye Color	Percent
Brown Eyes	50%
Blue Eyes	30%
Other Color	20%



© Center for Educational Assessment, UMass-Amherst

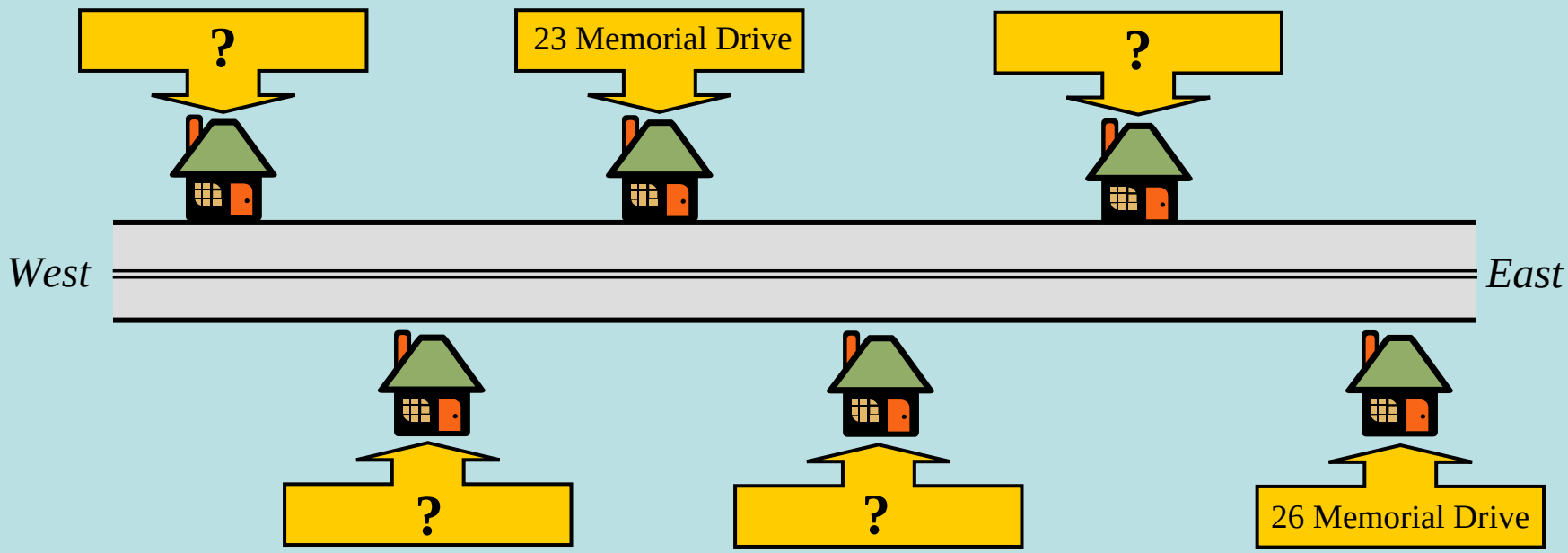
Select-and-Place

(Fitzgerald, 2001; Microsoft, 1999)

House numbers on Memorial Drive go west to east. Odd numbered houses are on one side of Memorial Drive, and even-numbered houses are on the other.

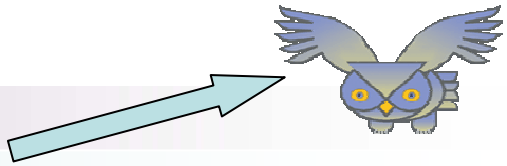
For each address in the table, determine which address corresponds to each house, and drag the address to the appropriate house.

22 Memorial Drive	24 Memorial Drive	21 Memorial Drive	25 Memorial Drive



© Center for Educational Assessment, UMass-Amherst

Delivery of New Tests



- On-line Web-Based Learning (OWL) is a project of UMass' Center for Computer-Based Instructional Technology (CCBIT)
  - Used by 1000s of students each year at UMass (> 20 departments); 1000s more at other unaffiliated campuses through licensing with textbook publishers
- CEA is partnering with CCBIT to develop tools for large-scale testing
  - i.e., item banking, adaptive algorithms, etc.
  - Delivery approach aligns with service mission of the University and the CEA

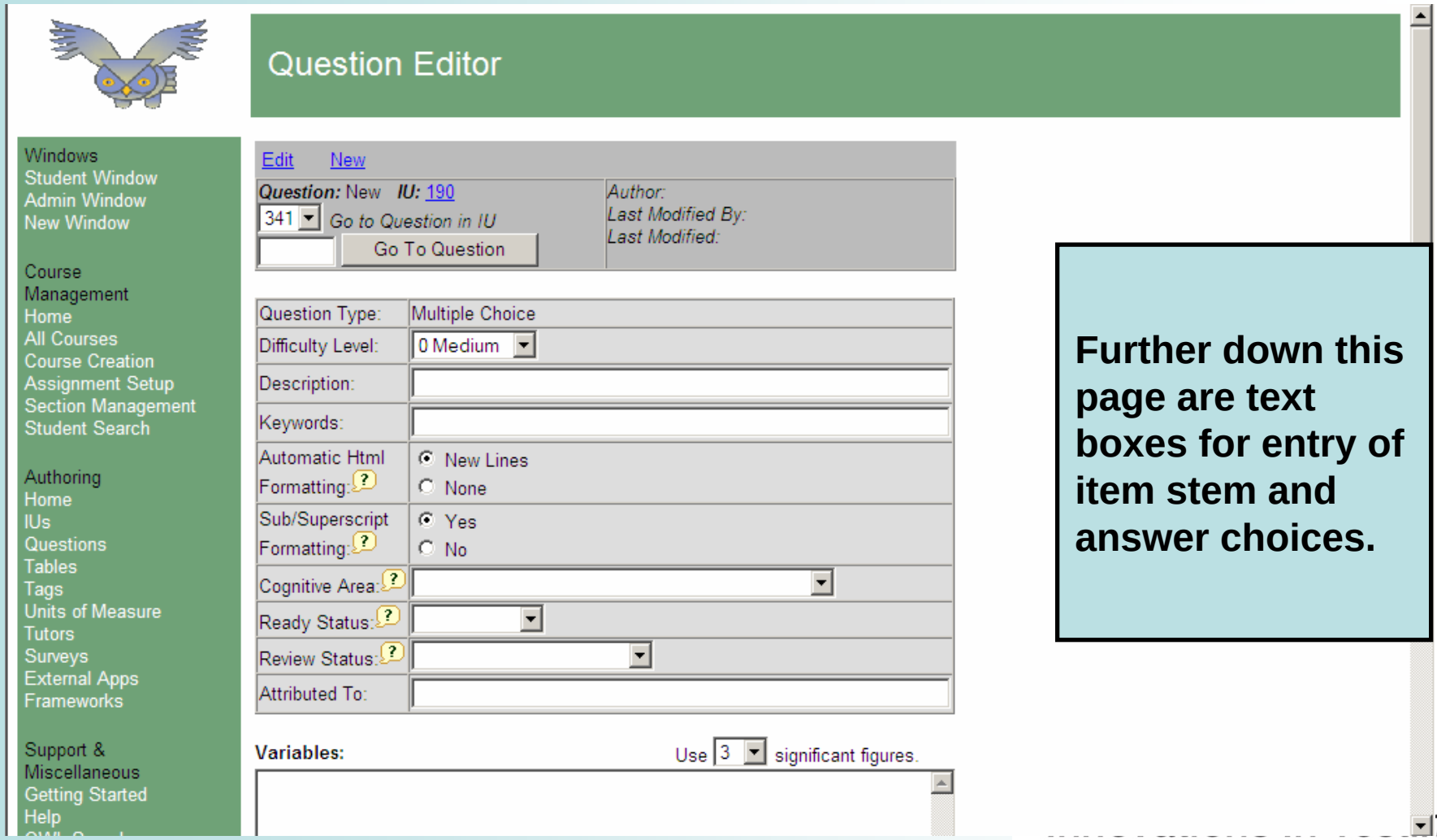
© Center for Educational Assessment, UMass-Amherst

Delivery of New Tests

- The OWL system includes
  - delivery/grading of electronic assignments;
  - course, homework and question authoring;
  - inclusion of course content materials;
  - course and student management tools;
  - and student progress tracking and grade reports
- OWL is web-based and accessible by password
- Students will log in to the system at ABE programs across the state to access their test

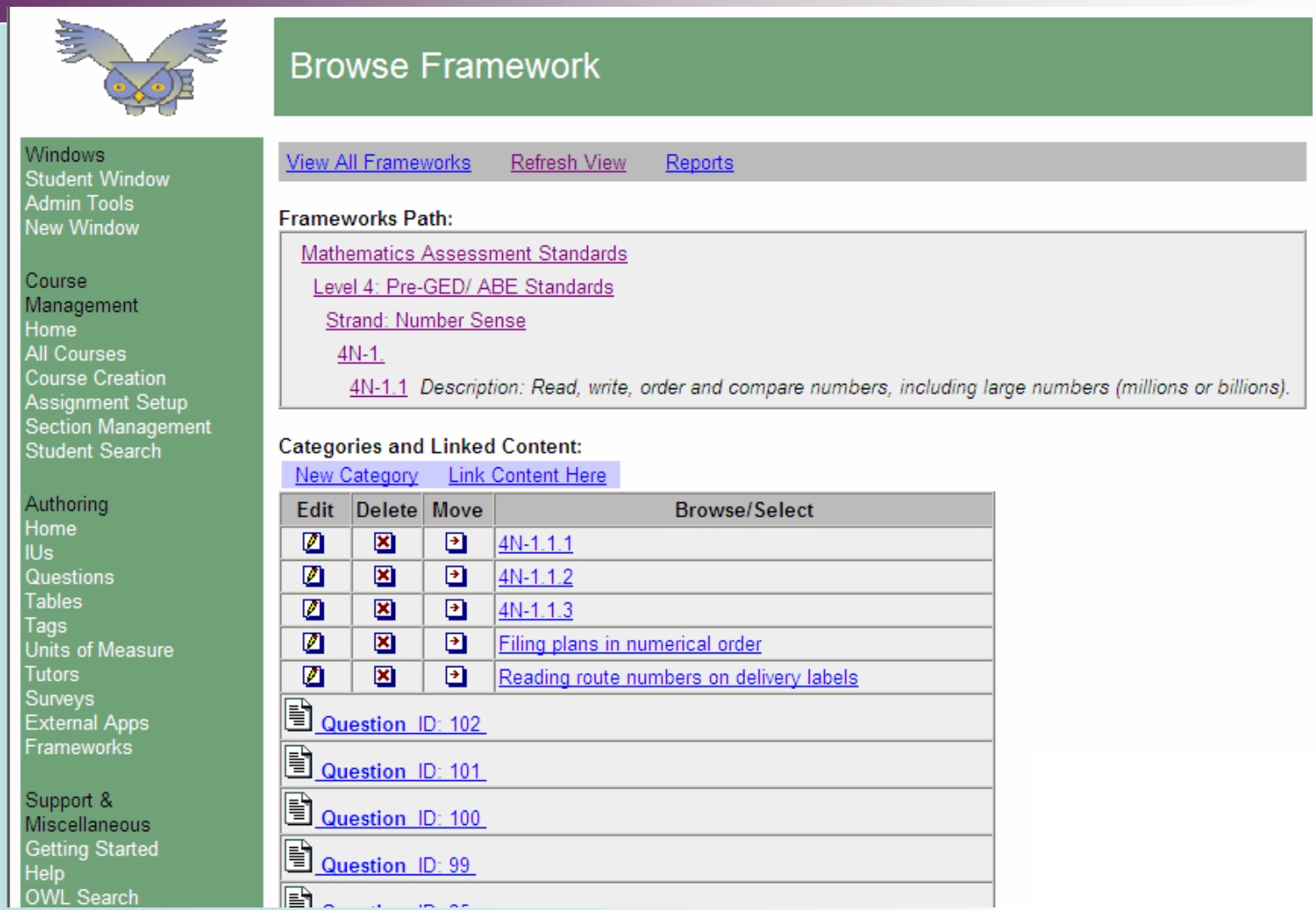
© Center for Educational Assessment, UMass-Amherst

Example: Web-based Question Editor



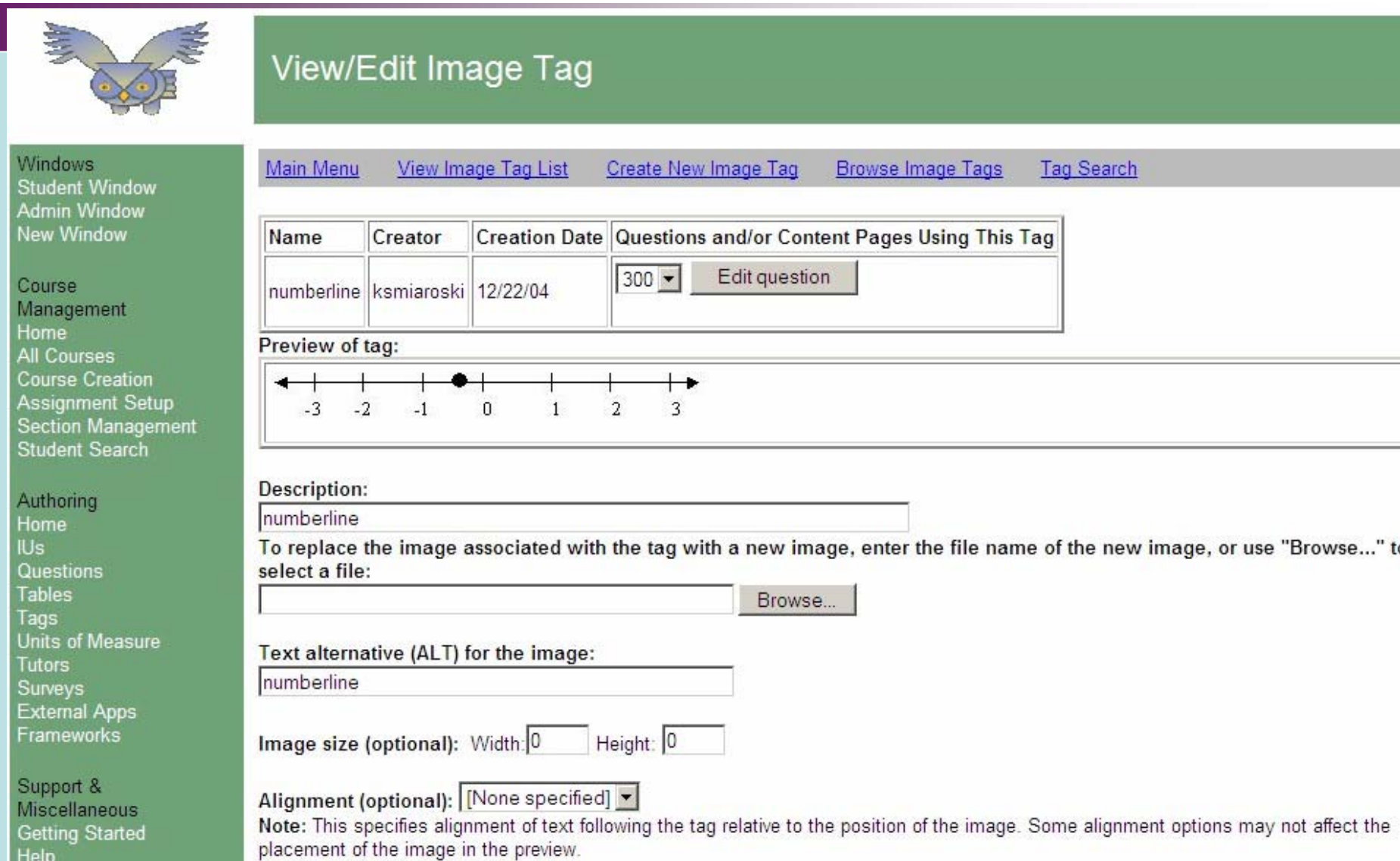
Further down this page are text boxes for entry of item stem and answer choices.

Example: Browse for Frameworks Links

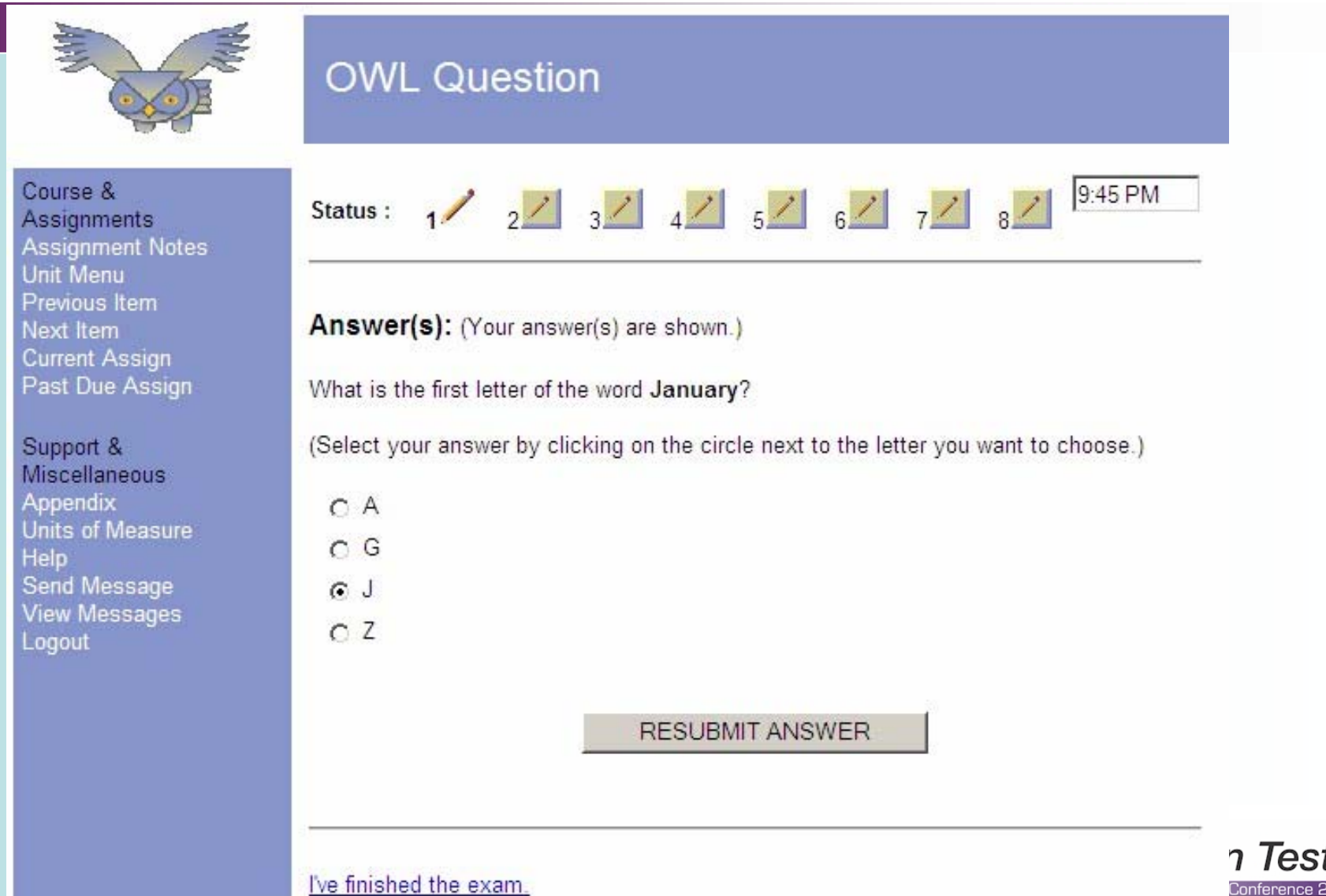


© Center for Educational Assessment, UMass-Amherst

Example: Adding Graphics to Items



Example: The Student Interface



Tutorial Development

- The testing population varies widely in computer familiarity, as found in a recent tryout study by Center staff
- The tutorial must serve two purposes:
  - Introduce basic browsing actions (click, scroll)
  - Assist students with navigating the OWL system
- An ABE teacher is working with Center staff to develop the tutorial to reflect simple language and use graphics where possible

© Center for Educational Assessment, UMass-Amherst

Questions?

email: azenisky@educ.umass.edu

for a copy of the slides:
http://people.umass.edu/azenisky