LUCID: A Web-Based Learning and Assessment System
Troy Wolfskill, Daniel Tofan, Amy Fong, and David Hanson
Department of Chemistry, Stony Brook SUNY
Stony Brook, NY 11794-3400

Introduction

The LUCID Project is developing a web-based system and materials to assist students and teachers in improving student learning outcomes. The system provides a rich set of guided-discovery activities, conceptual questions, exercises, and problems that can be used by students working individually or in teams. Team activities develop conceptual understanding and problem solving strategies, and use a peer review process to assess open-ended responses. Responses to exercises and problems can be collected in a variety of formats, including multiple choice, multiple response, fill-in, and drag and drop, as well as discipline-specific responses such as chemical reactions and Lewis structures. These responses then are analyzed to identify learning objectives that students have and have not achieved. By tracking student learning in real time by both topic and level of mastery, feedback can be provided to help both students and teachers improve learning. Since the learning activities that are part of LUCID have been described previously, this paper focuses on the assessment components.

The Need for Better Assessment of Student Learning

A recent summary of research into how people learn identified that effective learning environments need to be learner centered, knowledge centered, assessment centered, and community centered. Of these four, the least progress has been made in moving assessment to the center of Science, Technology, Engineering, and Mathematics (STEM) classrooms. STEM courses traditionally have a heavy dose of content knowledge. Classroom environments that engage students in the learning process and involve learning teams and learning communities are becoming more and more popular, and team problem-based learning has been a tradition in engineering, business, and medical programs. In contrast, assessment in STEM classrooms generally is limited to the assignment of grades that rank students according to their ability to perform on the quizzes and examinations that conclude courses and course units.

In order to improve student learning outcomes, students, teachers, and organizations need information on student learning that goes beyond grades. Students need to know the extent of their learning and receive feedback on how to improve their learning before they take examinations. Faculty need a clear measure of student learning so they can identify the materials and practices that most effectively enhance student achievement. Organizations need data regarding student achievement in order to use the available resources effectively to accomplish their goals. Students, faculty, and organizations need this information in real-time so they can adapt their processes, plans, materials, and resources to meet the immediate needs of learners.

While data regarding student learning is available from graded homework, quizzes, and examinations, resources for analyzing this data generally are not available. Students receive detailed information on the errors they make or a list of right and wrong answers, but they lack the skills needed to identify deficiencies in specific learning or problem solving processes. Too often they focus on their grade as a judgment of their ability, move on to new topics, and do not reflect on how to improve their understanding or skills or how to improve their learning process. Teachers may also have detailed information on student errors that would enable an item analysis of student performance, but the time, effort, and skills required for such an analysis are beyond those available to most faculty.
In the absence of meaningful analysis and feedback of assessment data, assessment practices can become incongruent with course goals. While essentially all teachers want their students to understand fundamental concepts and develop valuable skills, too often students can resort to algorithms, pattern matching, and memorization to solve homework and exam problems. Students identify the course goals from the exams not from the philosophy of the instructor, behave accordingly, and never develop conceptual understanding and problem-solving skills.

In order to provide better assessment of student learning, the following issues must be addressed.

- Quality measures of student learning must be available.
- Student responses to assessment items need to be stored in a computer for ready analysis.
- Strategies for analyzing student responses to produce meaningful information regarding student learning are needed.
- Easy-to-use and intuitive interfaces must be developed for students, teachers, and organizations to access these measures of student learning.

**Characteristics of a Learning Measure**

Essential characteristics of a measurement include a property, a scale, precision, and accuracy. We propose that the property to be measured in learning is competency. A competency is defined as a required ability to perform a task successfully. For example, the ability to identify the symbol of an element from its name is a competency. Different competencies can be associated with performance at different levels. These different learning levels comprise the scale. One way of defining the scale for measuring learning is to consider a students progress along a cognitive development scale or a scale of learning objectives. Blooms taxonomy of educational objectives provides an example of the latter. Using Blooms taxonomy, a competency can be characterized at the level of 1) information, 2) comprehension, 3) application, 4) analysis, 5) synthesis, or 6) evaluation. The precision with which a competency is measured can be characterized in two ways. The first regards the degree to which student performance with regard to such levels can be subdivided, e.g. can we assign a measure of 3.2 or only 3 to a particular student performance. The second regards the ability to resolve competencies. For example, a teacher might want to examine student learning at a low level of resolution, such as by chapter, or at a higher level of resolution, such as comparing the ability to balance redox reactions vs precipitation reactions. The accuracy of a learning measure is determined by a comparison to a standard. While standards for measuring learning are primitive in comparison to those for physical properties, one approach has been to score student responses to interview questions and compare those results with measures generated by other means.

**An Instrument Design for Measuring Learning**

With the characteristics of a learning measure in hand, we can consider the design of an instrument for measuring learning in analogy to a spectroscopic instrument. Figure 1 provides a schematic diagram of such an instrument. The essential components of this instrument are described below.
1. A database serves as a source of assessment items for probing student competencies. A single item may probe a single competency or multiple competencies. This database also contains information regarding the student and the course in which the assessment is being conducted. Items are parameterized to have variable quantities in order to support their reuse.

2. A student serves as the sample to be examined. Upon request of an item, the parameters in the item are replaced by actual values.

3. Student responses are submitted to a response analysis system that serves as the dispersion element for resolving the students performance with respect to the various competencies associated with the item. The result of this analysis is a record that shows which competencies were and were not exhibited by the students response. Each competency is associated with a learning level.

4. The competency record is stored in the database and feedback appropriate to the students performance is made available for display either immediately or at the conclusion of the assessment.

**The LUCID System**

LUCID, which stands for Learning and Understanding through Computer-based Interactive Discovery, is a web-based learning and assessment system that we are developing to implement the above design for measuring student learning.

**Principle Components of the System**

The principle components of LUCID are a course management system, a content management system, and an assessment management system.

**Course Management**

This component provides utilities for accessing the system, managing users, defining user roles, creating courses, enrolling students in courses, and creating and maintaining grade books. Users can be entered manually or uploaded from files. User roles include students, instructors, content developers, administrators, and software developers. Students and instructors may be enrolled in a course either manually or automatically by uploading suitably formatted files. Grade books are readily constructed by instructors using an easy-to-use interface. Grades may be entered manually, uploaded from files, or generated by the
software from assessment data. All users are provided with a username and password to access the system.

Content Management

This component provides utilities for authoring, browsing, selecting, and organizing learning resources for a course. Resources may be in any multimedia format supported by current browsers. Metadata for describing content enables users to search the system for resources by such things as discipline, educational level, and topic.

Learning Management

This component provides utilities for authoring assessment items and assessments, developing learning measures, identifying competencies, delivering assessments to students, analyzing student performance with respect to competencies, and reporting results of the analysis.

Contexts for Use

Student responses may be collected online or imported from computer-readable forms to enable the system to be used in a variety of contexts online or offline, proctored or unproctored, and by individuals or teams. The materials that we are developing primarily are intended for use in three specific contexts: 1) online team-based learning activities; 2) online individual assessments; and 3) offline individual assessments. Because of this flexibility, we plan to use the assessment system for in-class team-based learning activities on the web, for personalized individual online homework and quizzes, and for analyzing student responses on paper-based examinations.

Online Team-Based Learning Activities

The system includes utilities for a team of students to work on a computer in a computer classroom, and guided-inquiry learning activities are being developed for this purpose. These activities include an introduction to the competencies students are expected to develop and a guided inquiry aimed at developing conceptual understanding. The guided inquiry consists of an interactive model of a fundamental concept accompanied by a series of questions to guide exploration of the model, a set of skill development exercises aimed at systematically developing students' competencies in using and applying this concept, and problems requiring analysis and synthesis of newly developed competencies with previously introduced competencies. These activities have been described in more detail previously.

Responses to guided inquiry questions are open-ended and thus difficult to analyze by a computer. To assess student responses to such questions, a peer review process is employed. The system randomly selects one team's response to report for each question. As teams review these reports, they must either agree or disagree with the other team's response. These peer assessments are then compared to the teachers' assessment to determine each team's performance on the question. Peer assessments are summarized in a graph that shows the relative number of teams that agree or disagree with each report. The facilitator then can focus class discussion on disagreements to improve quality and achieve consensus.

Team responses to exercises and problems are analyzed by the response analysis system and instant feedback is provided. Utilities are provided for teams to document their reasoning in arriving at a response, and this documentation can be accessed for peer review. Competencies for the team are logged to the database separately from competencies associated with individuals. Following a classroom activity, each team member can access the activity along with the team responses and team competencies, though the responses cannot be changed.

Online Individual Assessments

Individual students may access the assessment system in five contexts: 1) personalized homework assignments; 2) personalized quizzes; 3) mastery learning activities; 4) confidence exams; and 5) proctored exams. Personalized refers to the manner in which parameters are varied. For example, a quiz item can
have parameters for concentrations, volumes, and chemical compounds so that students see a slightly different item in different contexts. These parameters can be varied by course semester, by section, by person, or each time an item is requested. When parameters are varied for the person (or personalized), the person sees the same parameters each time they view the item. When the parameters are varied upon a request for the item, the person sees different parameters each time they view the item.

A personalized homework assignment consists of a set of assessment items that are selected by the instructor for students to use in developing a range of competencies. Item parameters are set for the person so a student sees the same item each time the assessment is accessed. Instant feedback is provided for each response. Both learning measures and a grade are generated for the student to review, and items can be attempted repeated times until a correct answer is achieved.

A personalized quiz is similar to a personalized homework assignment with the exception that feedback is limited to noting whether a response is correct or incorrect. The instructor may set the number of attempts allowed. Access to such quizzes can be unrestricted or restricted to particular IP address to allow such quizzes to be limited to a proctored environment.

A mastery-learning activity steps a student through sets of competencies and learning levels. These sets are system-generated based on a students selection of competencies, learning levels, a particular range of difficulty, and the students performance on previous items. The parameters for each item change each time the item is accessed, thus a single item can be presented multiple times and have a different correct answer each time. A typical activity begins with items selected to assess competencies at a moderate level of difficulty. Successful completion of an item leads to a subsequent item with one of the following characteristics: 1) the same competency but a greater degree of difficulty; 2) a competency associated with the same topic but at a higher level of learning; 3) a new competency if the previous competency has been mastered at the desired level; or 4) mixed competencies requiring a higher degree of problem-solving skill. Unsuccessful completion of an item leads to a subsequent item with one of the following characteristics: 1) the same competency at the same level but a lesser degree of difficulty; 2) a competency associated with the same topic but at a lower level of learning; or 3) a previous competency if a failure with that competency is detected.

Confidence exams require that a student report not only answers but also their confidence that their answer is correct. In these exams, a fixed number of items is generated based on a students or teachers selection of a particular set of competencies, learning levels, and difficulty. Item parameters are personalized. No feedback is provided and only one attempt at each response is allowed. Each item is accompanied by an additional multiple-choice response for students to indicate their degree of confidence in the response. At the conclusion of the assessment, the student is provided with a grade and a learning measure that includes the confidence level for each competency. This coupled information helps the student identify issues that need to be addressed. For example, an incorrect answer with high confidence is especially noteworthy.

Proctored exams are similar to confidence exams except that access is restricted to particular IP addresses, all items are selected by the instructor, and no utilities are provided for indicating confidence.

**Offline Individual Assessments**

Assessments can be authored online and then printed for use by students in a proctored environment. Such assessments may be either personalized, or multiple copies may be generated in which only the order of questions and responses is varied. Student responses then can be collected offline using Scantron or other computer-readable forms.

**Supported Response Types**

To support the submission of student responses to the system, a variety of response gathering widgets are provided as listed in Table 1. While these include standard HTML form elements, additional widgets have been developed using JavaScript and Java Applets to support a broader range of responses. For example, styled text boxes allow students to enter molecular formulas and chemical reactions with symbols,
superscripts, and subscripts. The Lewis Structure Editing Tool allows students to draw a Lewis structure on a web page. Upon submission of the structure, a modified USMILES string is sent to the server for analysis.

Table 1. Supported Response Types

<table>
<thead>
<tr>
<th>Type of Question</th>
<th>Description</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>True/False and Multiple Choice</td>
<td>Select a single choice from a list of choices.</td>
<td>HTML radio buttons</td>
</tr>
<tr>
<td>Multiple Response</td>
<td>Select one or more choices from a list of choices.</td>
<td>HTML checkboxes</td>
</tr>
<tr>
<td>Short Answer</td>
<td>Input a word or short phrase.</td>
<td>HTML text input box</td>
</tr>
<tr>
<td>Numeric</td>
<td>Input a number</td>
<td>HTML text input box with JavaScript to assure a number is entered</td>
</tr>
<tr>
<td>Essay</td>
<td>Input one or more paragraphs of text.</td>
<td>HTML text area</td>
</tr>
<tr>
<td>Drag-and-Drop</td>
<td>Use the mouse to drag text images to predefined points on a page</td>
<td>JavaScript object</td>
</tr>
<tr>
<td>Image hotspot selection</td>
<td>Select a region of an image</td>
<td>JavaScript object</td>
</tr>
<tr>
<td>Select text</td>
<td>Select a portion of text.</td>
<td>JavaScript object</td>
</tr>
<tr>
<td>Order text and images</td>
<td>Drag text or images from one position in an ordered array to another.</td>
<td>JavaScript object</td>
</tr>
<tr>
<td>Styled text</td>
<td>Input a short answer that involves symbols, subscripts, or superscripts</td>
<td>Applet</td>
</tr>
<tr>
<td>Lewis structures</td>
<td>Draw a Lewis structure</td>
<td>Applet</td>
</tr>
<tr>
<td>Graphs</td>
<td>Sketch a graph</td>
<td>Applet</td>
</tr>
<tr>
<td>Drawings</td>
<td>Construct a drawing</td>
<td>Applet</td>
</tr>
</tbody>
</table>

A Preliminary Model for Measuring Learning

A scale for measuring learning allows a numerical value to be associated with a competency that in some way indicates the level of mastery at which that competency is performed. For preliminary studies we are employing an adaptation of Blooms taxonomy of educational objectives as follows.

1. **Information Level** - characterized by memorization, able to recall and repeat pieces of information, and identify information that is relevant.

2. **Algorithmic Application Level** - characterized by the ability to mimic, implement instructions, and to use memorized information in familiar contexts.
3. **Comprehension/Conceptual Level** - characterized by the ability to visualize, rephrase, change representations, make connections, and provide explanations.

4. **Problem Solving Level** - characterized by the ability to use the material in new contexts (transference); to analyze problems and identify the information, algorithms, and understanding needed to solve them, to synthesize these components into a solution; and to evaluate the quality of the solution.

For many competencies, these levels can be subdivided into responses that are of low, moderately low, average, moderately high, or high levels of difficulty. For example, while the ability to balance a chemical reaction equation is a level two competency, students asked to balance \( \text{Mg} + \text{S} \rightarrow \text{MgS} \) require a competency that is at a significantly lower level than those asked to balance \( \text{C}_8\text{H}_{19}\text{OH} + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O} \).

**Competencies and Topical Taxonomies**

In order to define a sufficiently fine scale of competencies that is suitable for use by the response analysis system described below, several hundred competencies are required for a typical introductory chemistry course. For example, the response analysis system requires four competencies for a simple ideal gas problem, one each for the determination of pressure, volume, temperature, and moles.

In order to simplify the authoring of assessment items, as well as to provide varying degrees of resolution for examining learning, disciplinary taxonomies are used to group related competencies with topics in a hierarchical structure. Figure 2 illustrates a portion of a taxonomy for General Chemistry that includes the above ideal gas law competencies. Each competency is additionally categorized according to the level of mastery it requires.

**Figure 2. An Example of the Use of a Topical Taxonomy for Organizing and Resolving Competencies.**

- Science
  - Biology
  - ...  
  - Chemistry
    - General
    - ...  
    - Gases
      - Pressure
        - The Ideal Gas Law
          - Solve the ideal gas law for pressure given ...  
          - Solve the ideal gas law for volume given ...  
          - Solve the ideal gas law for moles given ...  
          - Solve the ideal gas law for temperature given ...  
          - ...  
          - ...  
  - ...  
- ...  

**Item Authoring and the Response Analysis System**

In authoring items for use by the system, the presentation of the item is developed and parameters included. Parameters may be either numerical, in which case they are generated over a range of numbers at a specified precision, or they may consist of a list of values, in which case the list can either be declared explicitly or referenced to a field in a database table. Utilities for students to enter responses are provided and may include any of the response types listed in Table 1. A set of anticipated response conditions are then identified and each is linked to the successful or unsuccessful performance of one or more
The development of questions at the information and algorithmic levels generally is straightforward as they typically require a single response that is generated in a single step or standard sequence of steps. For algorithmic responses requiring discipline-specific symbols or mathematical equations, we have developed a set of Java classes to analyze a student's response to determine which steps in the algorithm have been completed successfully or unsuccessfully. For example, our chemical reaction equation classes compare a student response to the correct response to identify whether the appropriate symbols for the elements, reactant and product species, charges, states, and mass and charge balance are present. Our equation classes, which employ numbers with units, can be used to determine whether a student has rearranged the equation and converted units appropriately. As students respond to questions during formative assessments, they receive feedback on those steps they have completed successfully.

Differentiating learning at the conceptual and algorithmic levels has been discussed in the literature, and conceptual understanding can be assessed through single or coupled questions that have certain characteristics. Such questions have been used previously in the physics Force Concept Inventory, Mazurs ConcepTests, and the ACS General Chemistry Conceptual Examination. We have analyzed these question sets and identified the following nine characteristics of such concept questions. These characteristics allow us to readily identify, develop, and code conceptual questions. Responses to such questions are typically straightforward and can be readily analyzed.

- Ask questions requiring qualitative answers about situations described by equations. e.g. in Chemistry PV=nRT.
- Give a situation to analyze and predict the result or arrive at a qualitative conclusion.
- Change, interpret, or arrive at conclusions from a representation. (microscopic, macroscopic, pictorial, graphical, equation, tabular, verbal).
- Use equations with proportions rather than numbers so that plug and chug isn't possible.
- Take a concept and require that it be used in different contexts to reach conclusions.
- Ask a question plus a follow-up asking for a reason or match a statement and a reason.
- Complete a statement to make it valid or explain why it is valid.
- Ask a question with a response that requires making connections between two or more concepts.
- Identify which statements are correct or incorrect.

Problem-solving questions are straightforward to identify and develop and typically lead to a single short answer even though students require a range of skills to solve the problem. Some characteristics of questions that represent problems for students are given below.

- Require transference of prior learning to new contexts.
- Require synthesis of material from different contexts.
- Require analysis to identify information provided, information needed, algorithms and concepts required to connect these, followed by synthesis into a solution.
- Require developing a model, making estimates, making assumptions.
- Information is missing.
- Extraneous information is present.
- Has multiple parts.
- The route to the solution and the relevant concepts are not revealed by the question.

**Analyzing Responses at the Information, Algorithmic, and Conceptual Levels**

Anticipated response conditions may be generated in a variety of ways, depending on the nature of the response.

**Items Lacking Parameters.** For items that are not parameterized, anticipated responses are determined and linked to competency performance. For example, the following item can be used to assess a students competence in molarity at the information level

Which of the following are the units for molarity?
a) g/L   b) g/mL   c) mol/L   d) mol/mL

If a student submits mol/L, a successful performance of this competence is recorded in the database, otherwise an unsuccessful performance is recorded. The same competence could be assessed with a fill-in-the-blank response as follows.

The units of molarity are _________.

Anticipated responses that successfully meet the competence would need to include mol/L, mole/liter, moles per liter, and so forth.

The following item could be used in helping to assess students competence with molarity at a conceptual level.

When half of a solution of a known molarity is removed, the molarity of the remaining half: a) is halved; b) remains the same; c) is doubled.

**Numerical Parameters.** For responses developed from numerical parameters, anticipated equations that students might use are determined and linked to competency performance. A submitted response value is then compared to the list of anticipated values and linked to competency performance. If the submitted value is not found in the list, it is assumed that none of the associated competencies have been met. For example, a students competence with molarity and unit conversion at the algorithmic level might be assessed with the following parameterized item.

The moles of X in V milliliters of an M molar solution is _____.

Anticipated equations might include the following.

\[ n = \frac{M\times V}{1000} \quad \text{both competencies are met} \]
\[ n = \frac{M}{V} \quad \text{molarity competency is met but unit conversion competency is not} \]

In either case the values for n are generated dynamically by the system. Investigations into dynamically generating such permutations of equations have been explored and appear promising.

**Table-Based Parameters.** For table-based parameters, a submitted response generally is compared with the single response recovered from the table and linked to the competency performance. For example, the information level competency of knowing symbols for the elements could be assessed with the following item.

The symbol for ElementName is _____.

If ElementName is sodium, the lookup yields Na. If this is the response, the competency is logged as successful, otherwise not.

**Analyzing Responses at the Problem Solving Level**

The challenge with analyzing responses to multi-step problems arises in tracking the students problem-solving process and identifying how it can be improved. We have experimented with two methods for tracking such processes and both have provided satisfactory results.

Our first approach derives from an analysis of problem solutions that have a single answer resulting from the synthesis of several algorithmic steps. This technique identifies whether a student can take previously learned information and algorithms and transfer this knowledge into a problem-solving situation. A simple example is provided by a question that calls for a student to write a balanced equation for the reaction of sulfuric acid with sodium hydroxide. A correct response requires the student to synthesize previously developed skills in chemical nomenclature and formulas and balancing reaction equations with newly developed skills in predicting reaction products. Using our chemical reaction classes, which can identify
what has been done correctly and incorrectly, the correctness of the students response with regard to each of these competencies can be determined. In this example, competencies for chemical nomenclature and reaction equations are recorded at the problem-solving level because the student needed to transfer these skills into a new context and synthesize them into a solution. Prediction of the products is coded at the algorithmic level because that is the current topic and context. This approach can be employed with quantitative problems as well by generating permutations of possible answers based on common errors.

A second approach to analyzing responses at the problem-solving level provides two follow-up steps to determine the students ability to analyze the problem, employ prerequisite knowledge, and synthesize a solution. The first step has the student identify the competencies needed to solve the problem. If these competencies are consistent with the required competencies, a successful competency for analyzing problems is recorded. The second step has the student work through the sequence of lower-level competencies that lead to the solution. Failure in any one step identifies a failure in prerequisite competencies. Successful completion of both the analysis and series of steps implies a failure in synthesis. We are currently investigating ways to automate this process in order to simplify the authoring of the response analysis portion of items.

Reports of Learning Measures

Following an assessment, students receive reports of those competencies they successfully and unsuccessfully performed. The design of a more graphical measure is being developed that bears some resemblance to the measure shown in Figure 1 but includes reports for topics and multiple competencies. These graphical measures will have full support for increasing or decreasing the resolution by topic and level of learning.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grants DUE-9950612 and DUE-0127650. We gratefully acknowledge the National Science Foundation for support of this project along with Stony Brooks Department of Chemistry and Center for Excellence in Learning and Teaching. We also acknowledge Dan Apple of Pacific Crest as a contributor to the development of ideas. Finally, we would like to acknowledge the many undergraduates who have contributed to this project through software development, particularly Mo Chiu and Alfred Tsang.

Copyright 2003 by David M. Hanson, all rights reserved.

Bibliography


16. I.D. Eubanks, ACS Division of Chemical Education: *General Chemistry (Conceptual)*, 1996, Clemson University, Clemson, SC.