Thinking Ahead: the FIRST Assessment Database

Diane Ebert-May, Everett Weber, Mark Urban-Lurain, Ryan McFall, Matt Jones

Supported by the National Science Foundation DUE 618501

Student assessment data are collected in many different formats and at a variety of intervals within a course. Faculty also accumulate assessment data from the same course taught over successive years. Increasing numbers of science faculty want to analyze assessments in more detail and "do something" with the data from their courses. Results from their assessments can guide subsequent decisions about changing individual course designs and implementation strategies. In addition, if faculty had more options to compare common assessment results from within and outside their institutions, the results could influence even larger-scale changes in undergraduate biology curricula.

In order to facilitate this process of change, we are building a database to contain not only assessments but also individual student responses to those assessments. Metadata (data about the data) includes information about the institutional setting, classroom environment, and assessments. Faculty can use the metadata to find instruments to assess specific student learning goals and compare student responses from their own previous courses or courses taught elsewhere. The database simplifies data collection for individuals interested in cross-institutional research or longitudinal studies of their own course. The database is under construction now and a prototype will be online in January 2009. A major goal of this project is development of a flexible and effective metadata standard that allows users to find assessments and student responses to use for classroom studies and cross-institutional research.

Metadata describes and documents assessments. For example, what are the assessments? Where were the assessments conducted (i.e., institution)? When were the assessments administered relative to instruction? Why were the faculty members interested in collecting the data (i.e., summative assessment, validation of an instrument, effectiveness of web-based activities)? How were assessments delivered and student responses evaluated (e.g., rubric)? The answers to these questions are formalized in metadata so that other researchers can find data relevant to their research questions. We constructed a preliminary list of metadata for your review (Appendix 1), and request your input in determining what metadata are required to answer your questions about teaching and learning.

Searching the Database

The design of the database is based on the assumption that faculty will access the database to support their course development and research questions. The following examples illustrate two possible questions and the metadata required to find relevant datasets.

Question 1: An instructor wants to determine student learning gains about a topic (e.g., evolution) in response to a teaching innovation in a large enrollment course. What assessment tools are available?

The database will provide information about 1) instruments that are specific to the concepts taught by the instructor, 2) student responses to the instruments, 3) others that used the teaching innovation, and 4) specifics about the courses that use the instruments and teaching innovations.

Question 2: How do faculty determine if an assessment instrument is reliable and valid?
In this case, faculty are interested in testing the reliability and validity of an instrument. Ideally, reliability is assessed on the range of students and instructional environments in which the instrument is used. Instruments are tested across multiple institutions and include underrepresented students and those from a range of socioeconomic backgrounds. Both construct and criterion validity are determined with comparable instruments and interview data (Cronbach and Meehl 1955).

The database provides a means to include data from a range of institutions that serve a broad spectrum of students. Assessment questions from the instrument under validation are linked to questions or question clusters in another instrument to test validity. Faculty could use the database to collect and manage the student response data during the process of validation. For example, data from existing instruments such as the Concept Inventory on Natural Selection (CINS; Anderson et al 2002) are already included in the database. Faculty can use these data to further test the reliability and validity of the CINS instrument with a broader population of students.

The relevant metadata that faculty searching the database would need for both questions above include, but are not limited to:

a. Institution type (Carnegie classification)
b. Topic of course
c. Course enrollment
d. Course format (i.e., traditional lecture, on-line, hybrid, studio classroom)
e. Context of course (i.e., students' year in school, prerequisites for course, proportion of majors to non-majors)
f. Instructional strategies (e.g., cooperative learning, lecture, student-centered activities)
g. Timing of instruction (i.e., 1st week or 10th week of term)
h. Assessments
    i) Concepts tested
    ii) Publications (i.e., original instrument; studies that used the instrument)
    iii) Format of assessment (e.g., multiple choice, extended response)
    iv) Timing of assessment (pre-test, post-test)
    v) Assessment instrument
    vi) Rubric

Linking Metadata to Assessments
Individual courses, assignments, and questions can have multiple metadata entries. Suppose a research group is interested in evaluating the cognitive level of biology exam questions and rates individual questions with Bloom's Taxonomy. Three researchers rate all of the questions; each question then has three Bloom's ratings. Each rating is identified with the person that rated the question and the date the rating was uploaded. The researchers also rated each question for concept area (i.e., carbon cycle, evolution, and molecular genetics). A requirement of the metadata is that each rating (or classification) has a description, identification of the person that entered the description (for inter-rater reliability), and the date. The metadata provides a means

to document data in a searchable format. This helps ensure that changes in methodology, raters and question cluster alignment are properly documented so data can be monitored over time.

Data Submittal and Release

Student responses within the public portion of the FIRST Assessment Database are de-identified and institution identification, other than institution type, is protected information. There are two levels of protection integrated into the data submittal process. First, faculty upload data to a secure database with software that replaces student names and associated identifiers (e.g., PIDs) with a unique student number in each course (Figure 1). This allows researchers to link student responses from each individual's pre-test to post-test. Second, project personnel remove any demographic data and personal identifiers that were not removed by the software prior to releasing the assessment data to the public database.

To protect the identity of faculty and institutions, each faculty member and institution is assigned an Assessment Database ID number. If more information about a dataset is of particular interest to a researcher, project personnel will send a message from the inquiring investigator to the faculty member responsible for the dataset. The inquiring faculty member will not have contact information until s/he is contacted by the faculty member responsible for the dataset.

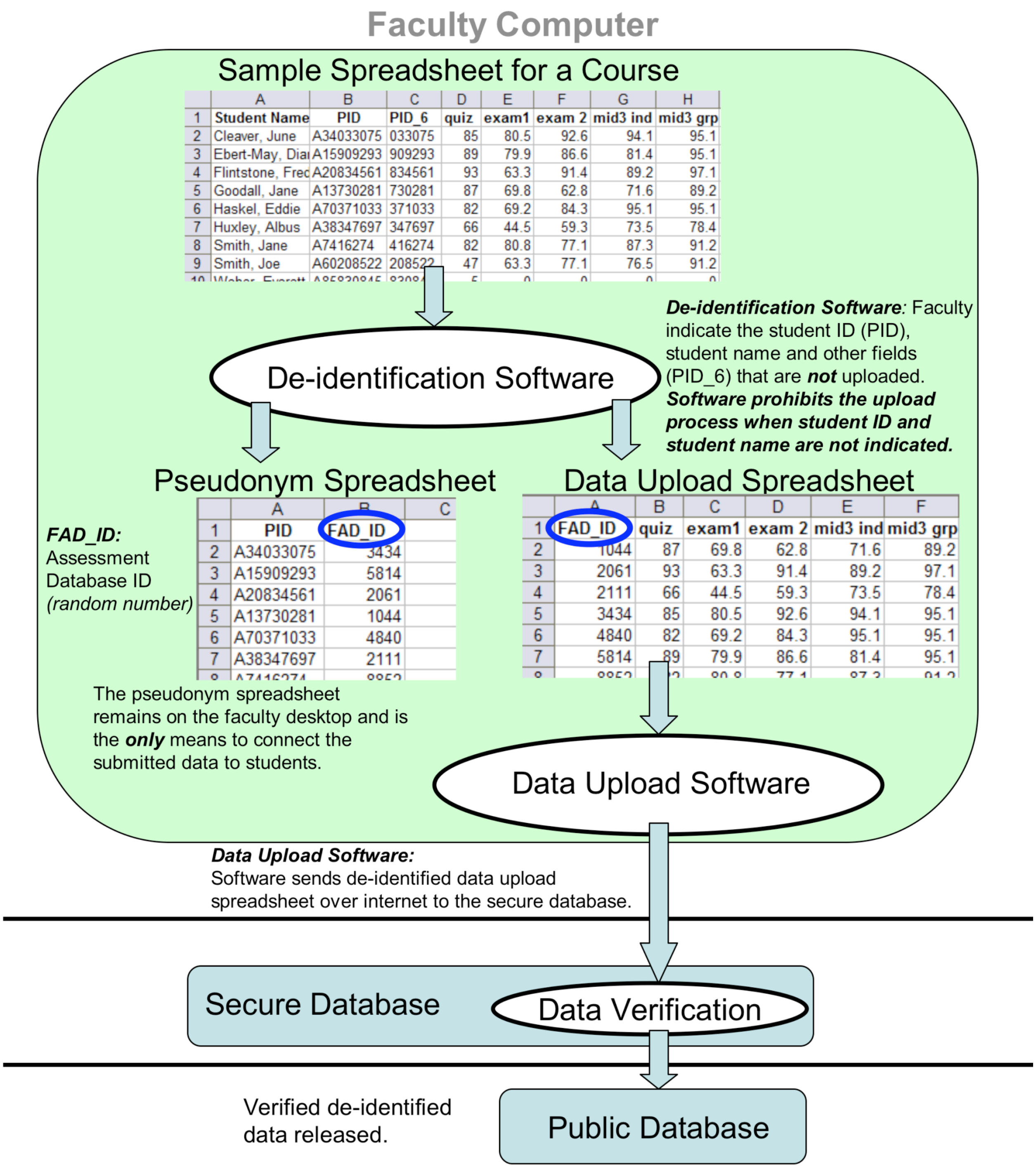
The Assessment Database, including the design for data submittal and release, is approved by the Michigan State University Institutional Review Board.

FIRST Assessment Database Development Team

Diane Ebert-May is the PI of this project and the team members include *Everett (Rett) Weber,* who is a postdoctoral research associate with Diane in the Plant Biology Department at Michigan State University. Rett contributes expertise in database development and built the pilot database for the project. *Mark Urban-Lurain* is Director of Instructional Technology Research and Development in the College of Natural Science at MSU and oversees the database management and software development. *Matt Jones* is from the National Center for Ecological Analysis and Synthesis (NCEAS) at UC Santa Barbara and will help develop educational metadata, build the 'back-end' of the database (i.e., what users never see), and adapt the existing NCEAS software to the database. *Ryan McFall* is an Associate Professor in Computing Science at Hope College, MI. Ryan will develop new software to create a user-friendly interface for the database. Collaborating with Ryan are NSF-funded REU students at Hope College.

AAAS and BiosciEdNet (BEN; Yolanda George) will assume all technical dimensions and oversight functions of the Assessment Database upon completion.

Figure 1. FIRST Assessment Database data submittal and release. When faculty upload their data, software replaces each student's identifiers (PIDs) with a unique random number (variable name: FAD_ID). As the data are uploaded to a secure database, software provides a spreadsheet linking the FAD_IDs to student IDs. Project personnel remove any demographic data and personal identifiers that were not removed by the software prior to releasing the assessment data to the public database.



National Metadata Committee

The Metadata Committee will oversee and contribute to development of the educational metadata standards, the minimum data requirements, and dataset quality criteria. Their input will significantly influence the structure and function of the database. The members include:

Carol Brewer, Professor, Division of Biological Sciences; University of Montana. Co-PI NEON project.

Malcolm Campbell, Associate Professor, Biology Department, Davidson College. Malcolm is a member of the ASCB Education Committee, former editor of LSE-Cell Biology Education.

Donna Sundre, Professor, Psychology and Executive Director, Center for Assessment and Research Studies, James Madison University.

Sam Donovan, Assistant Professor, Instruction and Learning, School of Education, University of Pittsburgh.

Gregor Novak, Professor of Physics, Indiana University. Gregor is the PI of JiTTDL – Just in Time Teaching Digital Library.

Nancy S. Shapiro, Associate Vice Chancellor, Academic Affairs, University System of Maryland.

Laurie Anderson, Assistant Professor, Department of Botany and Microbiology, Ohio Wesleyan University.

Leesa Sward, Professor, Department of Biology and Environmental Science, Valencia Community College. Orlando, Florida.

Terry Derting, Professor, Department of Biology, Murray State University.

Janet Batzli, Associate Director of Biocore, University of Wisconsin, Madison.

Request for CI Workshop Participants

In order to maximize the potential of the database, we request that you share with us your research questions about teaching and learning. Associated with each question, what kind of information would you like to find on the database? Consider the concept inventories in the context of the preliminary metadata criteria. What criteria work? What additional criteria should we add? This is a critical step in building metadata standards that are meaningful and useful to the biology education community. Your input is appreciated and valued.

Literature Cited

Anderson D, Fisher K, Norman G. 2002. Development and evaluation of the conceptual inventory of natural selection. Journal of Research in Science Teaching 39(10):952-978.

Cronbach LJ and Meehl PE. 1955. Construct validity in psychological tests. Psychological Bulletin. 52(4):281-302.

Appendix 1. Metadata Criteria – work in progress

Metadata Categories

These categories include descriptions of the place, time, instruction, instrument and student responses. This information is unlikely to change based upon interpretation of the submitted data. The data are arranged hierarchically from information about a course to descriptions of the question formats. Categories also include the student responses and the value assigned to a given student response (i.e. correct, incorrect, and/or specific misconceptions). In addition, categories such as taxonomies of understanding and concept categories of biology are used to further classify assessment items. Initially, faculty will likely search the database using concept categories to find instruments, questions and where the assessments were used.

Course Data

1) Institution
2) Department
3) Course name
4) Course number (include section)
5) Course size
6) Target audience
   a. Undergraduate/graduate
   b. Majors/non-majors
   c. Number of students in each class category (non-matriculating, freshman, sophomore, junior, senior, graduate)
7) Instructor name
8) Brief description of course
9) IRB approval for data collection (Yes/No)
10) Date(s) of instruction
11) First day of instruction in semester for topic (each lecture day counts as a day of instruction)
12) Syllabus
13) Course notes, activities and handouts for relevant assessments

Assessment Data

1) Assessment name
2) Assessment ID (links to questions, rubrics and student answers)
3) Instructions to class
   a. Verbal
   b. written
4) Was assessment group work?(yes/no)
5) Date assessment assigned
6) Date assessment due
7) Timing relative to instruction
   a. Before (pretest)
   b. During (in-class, mid lecture)
   c. After (posttest, can be after date of instruction)
8) Assessment format
   a. Clicker (Personal Response system)
   b. Written in-class
   c. Bubble sheet
   d. Online (provide link, if possible)

9) Credit
    a. None
    b. Points
    c. Percent of total grade
10) Graded for correctness
11) Electronic version of assessment (upload document(s))

Individual Assessment Questions

1) Assessment ID (links to assessment information)
2) Question ID (links to student answers and rubric/key)
3) Format
    a. Essay
    b. MC
    c. Matching
4) Credit (pts)
5) Question
    a. Text of question
    b. Options/sub parts
6) Question order

Rubric/Key

1) Assessment ID (links to assessment information)
2) Question ID (links to student answers and rubric/key)
3) Rubric code or correct answer
4) Description of code

Student Answers

1) Assessment ID (links to assessment information)
2) Question ID (links to student answers and rubric/key)
3) Student ID (random number unique to student)
4) Answer or rubric code (if multiple codes, separate by commas)

Categories of Learning

1) Bloom's Taxonomy (Bloom BS. 1956 Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.)
    a. Bloom 1956 (from Norman EG. 1970. Stating Behavioral Objectives for Classroom Instruction. New York :The Macmillan Company)
        i. 1 – Knowledge
        ii. 2 – Comprehension
        iii. 3 – Application
        iv. 4 – Analysis
        v. 5 – Synthesis
        vi. 6 - Evaluation
    b. Bloom Consolidated (Adapted by D. Sibley from Bloom BS. 1956 Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc)
        i. 1 – Knowledge and comprehension
        ii. 2 – Application and analysis
        iii. 3 - Synthesis and evaluation

Concept Categories of Biology

1) Evolution Concepts – includes the concepts of natural selection, and evolution.
    a. Anderson's 10 concepts (based upon Anderson et al 2002, misconceptions for each concept could be another related set of metadata)
        i. Biotic potential - All species have such great potential fertility that their population size would increase exponentially if all individuals that are born would again reproduce successfully.
        ii. Population stability - Most populations are normally stable in size except for seasonal fluctuations.
        iii. Natural resources are limited - Nutrients, water, oxygen, etc. necessary for living organisms are limited in supply at any given time.
        iv. Limited survival - Production of more individuals than the environment can support leads to a struggle for existence among individuals of a population, with only a fraction surviving each generation.
        v. Variation within a population - Individuals of a population vary extensively in their characteristics.
        vi. Variation inheritable - Much variation is heritable.
        vii. Differential survival - Survival in the struggle for existence is not random, but depends in part on the hereditary constitution of the surviving individuals. Those individuals whose surviving characteristics fit them best to their environment are likely to leave more offspring than less fit individuals.
        viii. Change in a population - The unequal ability of individuals to survive and reproduce will lead to gradual change in a population, with the proportion of individuals with favorable characteristics accumulating over the generations
        ix. Origin of species - An isolated population may change so much over time that it becomes a new species.
        x. Origin of variation - Random mutations and sexual reproduction produce variations; while many are harmful or of no consequence, a few are beneficial in some environments.
        xi. Biotic potential - All species have such great.
    c. Dino/Plant rubric (unpublished rubric by Ebert-May and Linton)
        i. P – change in individual (incorrect)/ change in population (correct)
        ii. C – need to change (incorrect) / change due to change in genes (correct)
        iii. V – All member of population are equally fit (incorrect)/ individuals within a population have varying levels of fitness (correct)
        iv. G – traits acquired during a lifetime are passed on (incorrect)/ genetic traits help the individual to survive and reproduce (correct) (partial if only mention survival or reproduction)
    d. Sexual selection – questions testing for understanding of sexual selection (a form of natural selection in which a male or female is the selective agent through mating preference).
2) Carbon cycling questions – questions that address the cycling of carbon in the environment. This includes photosynthesis, respiration, decomposition, consumption, provided the processes are within an ecological framework.

Instruments

1) Concept Inventory of Natural Selection (CINS) – questions from CINS published in Anderson et al 2002.