Problem-Solving Skills Among Precollege Students in Clinical Immunology and Microbiology: Classifying Strategies with a Rubric and Artificial Neural Network Technology

SUSAN KANOWITH-KLEIN,* [1] MEL STAVE,[2] RON STEVENS,[1] AND ADRIAN M. CASILLAS[3]

Department of Microbiology and Immunology[1] and Department of Medicine,[3] UCLA School of Medicine, Los Angeles, California 90095; Ulysses S. Grant High School, Valley Glen, California 91401[2]

Educators emphasize the importance of problem solving that enables students to apply current knowledge and understanding in new ways to previously unencountered situations. Yet few methods are available to visualize and then assess such skills in a rapid and efficient way. Using a software system that can generate a picture (i.e., map) of students' strategies in solving problems, we investigated methods to classify problem-solving strategies of high school students who were studying infectious and noninfectious diseases. Using maps that indicated items students accessed to solve a software simulation as well as the sequence in which items were accessed, we developed a rubric to score the quality of the student performances and also applied artificial neural network technology to cluster student performances into groups of related strategies. Furthermore, we established that a relationship existed between the rubric and neural network results, suggesting that the quality of a problem-solving strategy could be predicted from the cluster of performances in which it was assigned by the network. Using artificial neural networks to assess students' problem-solving strategies has the potential to permit the investigation of the problem-solving performances of hundreds of students at a time and provide teachers with a valuable intervention tool capable of identifying content areas in which students have specific misunderstandings, gaps in learning, or misconceptions.

Precollege educators and their students have increasing access to educational technology and, as a result, are teaching and learning in new ways that model the world outside the classroom (5, 23). Using effective technology, science educators and their students can bridge the gap from thinking about science to thinking like scientists, moving to a more authentic view of science and science research (3, 6, 21). With computer-based technology that promotes understanding of science content and integrates the use of scientific concepts, students are recognizing that what they learn in their studies has relevance beyond the classroom. Consequently, high school students are improving their content knowledge, honing their thinking skills, and beginning to consider their education in terms of professional opportunities in the sciences, opportunities that may lead from initial interest in life sciences to undergraduate and advanced degrees in the biomedical sciences (10).

Problem-solving skills are as critical for advancement in education as they are in the sciences. In contrast to rote learning and repeating exercises according to an algorithm, students must learn to develop the skills necessary to apply new information in a different context. The skills needed for problem solving may come as second nature to researchers and clinicians who use them as an integral part of their everyday work. However, precollege students first need to learn and practice these skills that are critical for their education in a fast-paced technological society (7).

Central to teaching and learning is knowing when a student reaches a particular level of understanding and what that level qualitatively implies about the student's understanding. The fact that nearly 30% of undergraduate freshmen enroll in remedial education courses (http://nces.ed.gov/pubs98/condition98/c9828a01.html) suggests that assumptions relating level of education with understanding are misleading. To investigate the degree of student understanding, educators are exploring new approaches that provide insight into the students' thinking skills and take advantage of advances in technology that provide timely feedback to teachers and ultimately their students (6, 13, 17). An ideal approach would be to observe the step-wise progression that students take through a real-life problem in order to understand what content knowledge students understand and if they know how to apply it. Methods, such as think-aloud, pen and paper, observation, and interviews, do exist to ascertain this sequential process (1, 24), but they are time-consuming on the part of both educators and students and limiting in the number of students that can be assessed in a given time (4). However, software technology does exist to follow large numbers of students as they solve complex problems (17, 18, 21).

Through our experience using the Interactive MultiMedia Exercises (IMMEX) program, we have demonstrated techniques that provide relevant problem-solving exercises and the assessment of the strategies employed through a range of learning levels from grade school to undergraduate and graduate courses to medical licensure examination with emphasis in the overlapping areas of microbiology, immunology, chemistry, and molecular biology (9, 16). Detailed information may also be found at the IMMEX website (http://www.immex.ucla.edu/ProfDevelopment/CaseStudy/frank1.html). Our assessment approach includes the visual representation of a student's step-wise progression through a problem. Using such a tool to provide feedback for their students, teachers can intervene where students' strategies indicate a lack of understanding, gaps in learning, or naïve misconceptions (6, 8, http://www.immex.ucla.edu/ProfDevelopment/CaseStudy/frank1.html).

Over time students who use this technology to track and report problem solving develop new strategies that relate to the mastery of content knowledge and its appropriate use (10, http://www.immex.ucla.edu/ProfDevelopment/CaseStudy/frank1.html). One may expect an evolution of strategic approaches so that an individual first learns to solve problems from a new content area as a novice and ultimately as an expert (2). We have begun to analyze the patterns of performance by students through the use of artificial neural networks (ANNs) as nonparametric, pattern-recognition tools to identify and cluster similar strategies (3, 16). The use of ANNs is a method of simultaneously classifying hundreds—even thousands—of strategies. Furthermore, when problems are complex, students may respond with a problem-solving approach that is hidden and not at all obvious within an intricate series of steps. ANNs can recognize patterns of similarity even when they may be well hidden within the data (14). Studies have shown differences between student and clinician performances in problems dealing with infectious diseases (19, 20) and have generated performance classifications independent of scoring criteria (16).

Here we report on the problem-solving strategies that high school students use to solve software case simulations in the context of commonly known infectious and noninfectious diseases, a novel approach for the high school biological sciences curriculum. In addition, we report on methods to classify these strategies, methods that reveal facts, concepts, and correlations students do and do not understand. With such information, educators can provide feedback and specific interventions for the students toward increased understanding and enhanced problem-solving skills. Finally, we have adopted a neural network assessment approach toward analyzing performances in clinical immunology and microbiology topics appropriate for high school students.

MATERIALS AND METHODS

Interactive, problem-solving simulations. As a recipient of a John H. Wallace 1998 Summer Teacher Research Fellowship, sponsored by the American Association of Immunologists (AAI), high school teacher Mel Stave authored the problem set Antibody with IMMEX software (18). Additional details may be found at the AAI website (http://12.17.12.70/aai/committees/education/fellowship.htm, http://12.17.12.70/aai/committees/education/Curriculum/auntie_body.htm). Based on the understanding of microbiological and immunological concepts, Antibody consists of 10 parallel cases. Each case opens with a prologue detailing the problem to be solved: an intern at a teaching hospital must suggest a diagnosis for a clinic patient who is described in the prologue. Using information obtained through a menu-driven format, the students solving the problem have access to a variety of reference and test data items that might be used to solve the case. For example, within the problem space students can access results of Northern, Southern, and Western blots, enzyme-linked immunosorbent assays (ELISA), and skin prick tests (SPT, a physiologic test to determine presence or absence of a specific allergy). Once students think they have collected, interpreted, and synthesized enough data and other information from the menu items, they can try to solve the problem. When the students complete the problem, an epilogue or summary of the problem appears. The epilogue outlines approaches that students can take to solve the problem correctly and serves as a model for solving additional cases. In this study, students were presented a module of 10 cases consisting of clinical problems: four relating to allergies (food, shellfish; chemical, latex; pollen, weed; and animal, dust mite), two each to bacterial and viral infections (*Pneumococcus*, *Streptococcus*, herpes simplex virus, and hepatitis A, respectively), one to tumor growth (ovarian), and one to a healthy individual.

Detailed information about the Antibody problem set and a sample case to be solved can be found at the IMMEX website (http://www.immex.ucla.edu/HomeMenuItems.htm) click on the link to "IMMEX on the Web"

(http://www.immex.ucla.edu/MedicalHealth/AAI/AuntiebodyPoster.pdf). Educators can request use of the complete Antibody problem set for their students at http://day.immex.ucla.edu/IMXWeb/Staging/staging_request.htm.

Analysis of student performances on simulations. As students solve the problems, a database is generated to trace all the menu items accessed by the students as well as the order in which they accessed these items. This database is then used to recreate student strategies in the form of search-path maps, visual representations of the students' problem-solving approaches (21). A search-path map consists of a template displaying all the possible menu items that can be selected by a student. Lines connect the items and indicate the sequence of test items selected by the student. Examples of search-path maps are shown in Fig. 1 and 3. Additional details may be found at the IMMEX website (http://12.17.12.70/aai/committees/education/Curriculum/auntie_body.htm).

Classifying strategies with a rubric. To evaluate the quality of the students' problem-solving strategies, we prepared search-path maps of the student performances. These search-path maps demonstrated multiple approaches to solving the Antibody problem set. To characterize and order these strategies regardless of the students' point score, we developed a classification system, a rubric based on specific criteria and gradations of quality. The student performances themselves defined the rubric into five levels or major strategies (described in the Results and Fig. 1), thus resulting in no *a priori* classification of the performances. All the maps were then rated with the rubric by two groups of people: three content experts in the area of high school biology who used the Antibody cases for teaching and two university-based educators with expertise in immune-related diseases.

Neural network architecture and validation. Software based on artificial neural networks (ANNs) can build rich models of complex phenomena through a training and pattern-recognition process. Such networks have had significant, practical utility in solving classification problems with ill-defined categories, where the patterns are well-embedded within the data and where models of behavior are not well defined (11, 12, 14). We have previously described the specific methodology used in relation to problem-solving performances and have shown that ANNs can analyze patterns of performance that span broad levels of achievement (3, 16).

Thus, we took advantage of the ability of ANNs to cluster performance data based upon the performance itself. The unique tests or individual menu items selected by each student constituted the input pattern so that a distinctive set of inputs represented a student strategy to the neural network. The menu items selected were preprocessed, resulting in a binary representation (1 = test items selected, 0 = test items not selected). In the training process the inputs were repeatedly presented to the network. This presentation resulted in a grouping or clustering of the inputs onto a grid that was defined as a two-dimensional "output" space (5 by 5 grid). Thus, all the performances appeared on this output space in groups that related performances by similarity (3, 16). Additional information may be found at the IMMEX website (http://www.immex.ucla.edu/MedicalHealth/Ottawa/intro.htm).

Visualizing strategies clustered by the neural network. Once the ANN clustered student strategies, we used the IMMEX analysis software (18, 21) to generate a search-path map of the clustered performances. In contrast to a search-path map of a single performance (Fig. 1), these group search-path maps represented all performances, from 1 to 71, that were assigned to a single node. The thickness of any given line that connects two menu items is directly proportional to the number of students who selected the same sequence of menu items (Fig. 3A).

RESULTS

Student performances. A total of 470 student performances of the Antibody problem set were collected from 10 high school classes, representing five teachers in three Los Angeles Unified School District high schools during the academic year 1998–1999. The 10th–12th grade students who solved the multiple-case problem set were studying biology in regular, honors, and advanced placement (AP) biology classes.

Students solved 67% of all 10 cases in the Antibody problem set. For each case completed the students received a score, based on points deducted for each item selected from a total starting score of 1,000 points. No statistically significant differences ($p < 0.05$) were observed in scores among the cases, suggesting the cases were of equal difficulty. Although these data indicated that the students were able to solve all the cases within the problem set, they revealed little about the

quality of the students' approaches to solving the problems, that is, either the sequence of steps taken to solve the problem or the predominant menu items used.

Looking at search-path maps and assessing problem-solving strategies: performance analysis by rubric assessment. To evaluate the quality of the individual students' problem-solving strategies, we studied the individual student's search-path maps and developed a classification scheme (rubric) that reflected qualitative differences among the strategies. Five major classifications were apparent. The key determinants of the rubric classification focused on the use of menu item data and subsequent actions stemming from interpretation of the data. Thus, with the latex allergy case as an example, the best-solved strategies consisted of appropriate data acquisition that immediately led to the solution (rubric level 1; Fig. 1A). Less efficient, but solved strategies were those for which appropriate data were collected yet the solution was not immediately apparent to the student (rubric level 2; Fig. 1B). Even less effective strategies were ones in which students never accessed appropriate data and may have used roundabout methods of obtaining the correct solution (rubric level 3; Fig. 1C). Differences were also observed among unsolved performances. Some students failed to deduce the correct answer despite having obtained appropriate data (rubric level 4; Fig. 1D), whereas other students failed to obtain any relevant data and were given the poorest score (rubric level 5; Fig. 1E).

The development of a rubric that encompassed a diverse array of performance patterns was meaningful in that different levels of the rubric defined specific types of strategies based on a sequence of items selected. However, a major drawback of such a system is that a teacher must read and manually score each map. In the present case, five experts met to reach consensus on scoring the maps, a time-consuming process. Even if there were a way to overcome this obstacle, not all strategies fit the rubric. In such situations, general classification imposed by a rubric may encompass only a portion of the desired model being developed. Since the primary method of classifying strategies with the rubric resided in the patterns of test item selection, we then investigated the pattern-recognition capabilities of ANNs to automate and expand our strategy classification methods.

Artificial neural network analysis. After training unsupervised neural networks with the 317 solved performances, we used the same training data plus 153 unsolved performances (test data) to determine where major clusters of performances were located in the ANNs' output. Thus, we observed performance clusters at multiple locations or "nodes" (Fig. 2A), with 21 nodes containing clusters of 10 or more performances (Fig. 2B).

The cases dealing with allergic disease were most similar in terms of the strategic approach to solving the problem. We observed that for nodes 3, 6, 7, 15, 17, and 19 more than 75% of all performances at these nodes involved an allergy solution (Fig. 2B). By examining individual performances within each of these six nodes, we found that 100% of the allergy cases were solved within four of the nodes (6, 7, 15, and 17) while 75 - 78% of the allergy cases were solved in the two remaining nodes, respectively 3 and 19.

Since the percentage of solved allergy problems at these "allergy" nodes was high (75–100%), this excellent solve rate indicated that students were able to recognize the difference between an allergy and a nonallergy case and subsequently used an appropriate allergy-specific strategy. In fact, when we examined performances at individual nodes using search-path maps and item usage from the neural net results, distinct allergy-specific patterns or student strategies were evident, specifically at node 3 for food allergy (Fig. 3) and node 19 for pollen allergy (Fig. 4).

For example at node 3 (n = 26 performances; 14 dealing with food allergy, 12 dealing with other illness), all 14 food allergy cases were solved while few (2 of 12) of the remaining nonallergy cases were solved. Students whose performances were clustered at this node almost exclusively selected the results of an SPT for a specific food allergy (shellfish), and over half the students chose the results of an ELISA for a food allergy (Fig. 3). These choices represented an appropriate strategy to determine a food allergy but not an appropriate strategy for other cases. In fact, the 10 unsolved cases that included both nonallergy and other allergy cases received a rubric score of 5, the lowest score possible. Students with the latter performances either did not understand how to interpret the test results selected or were wildly guessing.

Next we investigated the relationship between the ANN-generated clusters, based on menu item selection, and the rubric classifications, based on the sequence of menu items selected. We found that the solved allergy cases at all allergy nodes had good rubric scores, the majority in the highest category. Most allergy cases (68 of 92 or 74%) received a score of 1

(appropriate selection of menu items and efficiency in solving the problem); 8 of 92 or 9%, a score of 2 (appropriate selections but more searching); and only 4 of 92 or 4%, a score of 3 (inappropriate selections but correct solution). In contrast, unsolved allergy cases were few: 1 of 92 or 1% received a score of 4 (appropriate selections), and 8 of 92 or 9%, a score of 5 (inappropriate selections). These results implied that the allergy performances clustered at the allergy nodes where students used appropriate tests in efficient ways to solve the cases and signified a positive relationship between the ANN and rubric results. We also determined by the Pearson chi-square test that the difference between the number of allergy cases solved at allergy nodes and nonallergy nodes was statistically significant ($p = 0.017$).

Predicting a strategy from the ANN cluster. To test the hypothesis that neural net clustering could predict a type of strategic approach or at least determine how the network was clustering performances, we created a series of mock performances (i.e., a series of steps from one menu item to another) to predict the nature of a given performance that causes it to fall within a specific cluster. This form of analysis was clearly retrospective in nature, since we had prior knowledge of the main features resulting in clustering at certain nodes. However, it served to illustrate that novel performance strategies (never encountered by the network during training) could be predicted. We used a variety of strategies to generate the mock performances. For example, we solved allergy cases by selecting (i) the ELISA data only, (ii) SPT data only, and (iii) ELISA data followed by other test results or information. Then we input the resulting performances into the trained neural network.

We found we could predict an output by providing the ANN with a sequence of tests we expected to appear at certain nodes. For three of the six major allergy nodes (nodes 3, 15, and 19), the ANN correctly matched the mock performances to the expected nodes where the percentage of allergy cases was greater than 75 (data not shown). As shown above, almost all of the student performances at node 3 took advantage of the SPT for food and more than half, the ELISA for food (Fig. 3B). The mock performance using the SPT for food fell at node 3, validating that clustering at a specific node predicted a particular strategy. In contrast, the mock performance using the SPT for pollen fell at node 19, a unique node containing only allergy cases that were predominantly the pollen allergy case. At this node 90% of the students selected the SPT for pollen (Fig. 4A). Finally, the mock performance using an ELISA first and then other data selections fell at node 15, the node at which student performances relied on multiple ELISAs. 100% of the students accessed ELISA for food; more than 90%, ELISA for dust mite; and about 65%, ELISA for both chemicals and pollen (Fig. 4B).

The remaining three major allergy nodes (6, 7, and 17) contained performances that were more complex. They included the same major item selections as at nodes 3, 15, or 19 in addition to 8–9 other menu items. For example, at node 6 the item selections chosen by 60% or more of the students included allergen and SPT information, allergist, SPTs for dust mite and food, pollen count, and ELISAs for dust mite, chemicals, and food (Fig. 4C). These multiple selections contrast performances at node 15 where selections by more than 60% of the students only included ELISAs for dust mite, chemicals, food, and pollen (Fig. 4B).

FIG. 1. Classification with a rubric: alternative approaches to solving the chemical (latex) allergy case. (A) Rubric score: 1: student accessed relevant data (circled) and moved directly to the solution. The search-path map is created by overlaying the student performance or sequence of student selections on the template and by clearing the menu items not selected. The line moves sequentially from the upper left-hand corner of the first menu item to the bottom-middle border of the next menu item. The different shades of the boxed menu items indicate these selections are from different content domains available in the main menu of the problem. Additional details may be found at the IMMEX website (http://www.immex.ucla.edu/HomeMenuItems.htm). (B) Rubric score 2: stduent accessed relevant data (circled) and strayed before solving the problem. (C) Rubric score 3: students accessed irrelevant data and still solved the problem. (D) Rubric score 4: stduent accessed relevant data (circled) but did not solve the problem. (E) Rubric score 5: student accessed irrelevant data and did not solve the problem.

FIG. 2. Major nodes. (A) Student performances clustered by the ANN were assigned to specific nodes. Nodes 3, 6, 7, 15, 17, and 19 were composed of at least 75% allergy cases. Nodes 18 and 24 contained a significant number of cases relating to viral diseases. Node 8, the largest node (71 performances), contained many unsolved, incomplete, and guessed (direct from start to solution) performances. (B) Nodes with 10 or more performances showing the ratio of allergy to nonallergy cases attempted at each node. Nodes with 75% or more allergy cases are indicated with an arrow.

FIG. 3. Student performances at node 3, search-path map vs. menu item usage. (A) A group search-path map showing 26 student performances clustered at node 3. The thickness of the lines indicates the relative number of students who selected the same sequence of menu item usage. The two most accessed menu items at this node (circled) were the SPT and ELISA for food. All 14 food allergy cases at this node were solved in contrast to only two of the remaining 12 cases. (B) More than 95% of the students selected the menu item with the SPT results for food, and more than 50% selected the menu item with the ELISA results for food. All students automatically selected item 1, the case scenario from where the problem begins.

FIG. 4. Menu item usage for student performances at Nodes 19, 15, and 6. (A) Menu item selection versus percentage of student performances at node 19. More than 90% of the students selected the menu item for pollen SPT results and about 65%, the menu item for pollen ELISA results. (B) Menu item selection versus percentage of student performances at node 15. Four menu items were selected by more than 60% of the students and all of them were ELISA results. More than 65% of the students selected menu items for ELISA results for both a chemical (latex) and pollen, while close to 95% selected the menu item for ELISA dust mite results and 100% for ELISA food results. (C) Menu item selection versus percentage of student performances at node 6. All students conferred with the allergist while 60% looked up "allergen" and "skin prick test" in the library. Close to 90% of the students selected menu items with SPT results for dust mite and food and with ELISAs for dust mite and food, while 80% accessed ELISA results for the chemical latex and for the pollen count.

DISCUSSION

We have shown that (i) a rubric can be developed for scoring problem-solving performances or strategies, (ii) ANNs can cluster similar performances or strategies, and (iii) a relationship exists between the rubric results and performances clustered by the ANN. The latter suggests that the quality of a problem-solving strategy can be predicted from the location of the clustered performances to which it is assigned. Inherent to the process of analyzing a strategy is the detection of certain patterns in the student's performance. When students solve real-life problems that require integration and application of knowledge, they demonstrate a variety of problem-solving strategies—ranging from weak to effective—that reflect their understanding of the task and its solution (15). Such a problem-based approach poses certain requirements for adequate assessment: (i) the need to visualize strategies in an efficient manner, (ii) the ability to classify the strategies according to a level of mastery and competence, and (iii) the time necessary to sort through many individual strategies.

We have addressed these requirements for assessment with search-path maps, a rubric, and the use of ANNs. As a qualitative measurement, search-path maps visualize a single student's or a group of students' strategies to solve a problem. Although relatively time consuming, rubric scoring of individual search-path maps provides a quantitative measure of the approach the students use to solve a problem. Furthermore, the rubric permits identification of the level of understanding a student has for content presented in the problem. In contrast, clustering the performances with ANNs provides a comparatively rapid method to separate different types of problem-solving strategies from a mixture of performances and simultaneously group similar performances in the same cluster.

In addition to allergy performances that were clustered at allergy-specific nodes by the ANN, we also identified nonallergy performances that clustered at specific nodes. For example, at node 18 where there were no allergy performances, 10 of 10 cases of viral illness were solved successfully and efficiently with three predominant menu items accessed—information for herpes simplex and hepatitis A and ELISA data for viruses (data not shown). Likewise, at node 24, 9 of 13 cases of viral illness were solved out of a total of 32 performances. Although only 41% of all the performances represented a viral case, more than 95% of the students selected the ELISA data for viruses (data not shown)—a major feature of performances at this node.

Interestingly, the students discovered approaches to solving the cases for bacterial and viral infections, approaches that were unexpected by their teachers. Using an anticipated approach, one group of students based their conclusions on the interpretation of scientific evidence, i.e., by analysis of ELISA or Western blot data. In contrast, a second group derived its answer from a clinical description of viral and bacterial diseases in the library. Although both approaches are valid and speak to the variety of successful methods students can follow to solve problems, the scientific approach can be generalized, being adaptable to many different situations in which ELISA or blot data is available. However, the use of clinical descriptions is limiting in its transfer to other problem situations where a specific clinical description may not be available.

With 71 student performances, node 8 was by far the largest node (Fig. 2) and contained a mixture of performances. For example, performances that fell at node 8 included the majority of incomplete performances, guesses (directly from start to solution with no menu items accessed), and performances with an enormous amount of search, indicating that these latter students were lost. In addition, mock performances in which there was considerable menu item selection fell at node 8 (data not shown). Moreover, node 8 contained the largest number of rubric scores that were 3's (18 out of 70 performances) (i.e.,inappropriate data accessed and the problem solved correctly) and 5's (18 out of 70 performances) (i.e., inappropriate data accessed and the problem not solved correctly.

Overall, unsolved performances did not fall at the major allergy nodes or at the viral nodes (noted above). Unsolved performances were characterized by the selection of an extensive and inefficient number of menu items. The common features of the major allergy nodes and viral nodes were that relatively few menu items were selected by most of the students, indicating that these strategies were deliberate.

Another important issue for consideration is how students frame a problem, i.e., what are the preconceived notions students have about a problem. For example, students whose performances landed at allergy node 3 and who did not solve the problem did not frame the problem correctly. In spite of conflicting evidence as more information was collected, 12 of 26 students continued to follow a food allergy strategy. If they had not, their performances would have landed at another node. Did these students previously solve an allergy case and assume that a similar strategy would work? With the analysis software the sequence of the student's performances can be investigated to follow the development of such inappropriate strategies.

For the Antibody cases described, this interactive software approach is valuable in identifying students who understand or do not understand basic concepts that relate to the diseases and diagnostic tests needed to identify each case. With search-path maps, viewed either one-on-one, ranked with a rubric, or clustered by nodes, teachers can provide appropriate classroom interventions for improving the students' content understanding and developing the students' problem-solving skills. In addition, they can use such information to follow student progress over time. For example, with student practice over time teachers can look for search-path maps that become more efficient in solving a problem, rubric scores that move from levels 4 and 5 to levels 2 and 1, and performance clustering at nodes that shift from guessing and disorganized searching to an efficient strategy.

The IMMEX Laboratory is pursuing methods to intervene during a student's problem-solving performance. The types of analyses presented in this paper may lead to real-time interventions during problem solving in later versions of the software application. In the meantime, each problem set is made up of multiple cases. Instructors can intervene at the conclusion of a student's performance and discuss that performance with the student before he or she moves on to the next case. Additionally, each case has an epilogue that provides the student with a means for self-intervention.

As an assessment tool, the rubric described in *Results* clearly defines differences in strategies between solved performances (rubric scores 1–3) and not-solved performances (rubric scores 4–5). Students who solved the problems with rubric scores 1–2 were able to apply content knowledge to demonstrate understanding of the concepts. The inference may be that students who score well with the rubric may likewise do well in the course. Although we do not yet have data to make such a statistical correlation, we are moving in this direction.

A correlation, however, does exist between the ability to solve many different problem sets and course grades. Data from three high school AP and regular science classes indicate that student performances (solved versus not-solved performances) across all problem sets completed during an academic year are moderately correlated with course grades (0.4–0.6) and highly significant ($p = 0.001$) (22). The moderate correlation is likely due to the fact that the software problem sets are measuring multiple types of skills and understanding in addition to content knowledge.

Since problem-solving scenarios are complex by nature, distinguishing a student's content knowledge from his or her understanding of how to apply content knowledge can be difficult. In addition, behavioral issues enter into play. For example, the student's level of confidence in knowing that he or she has enough information to solve the problem can influence strategy. Students with rubric score 2 already appear to have enough information to solve the problem but continue to collect more. They may not understand the information they already possess or simply may lack the confidence

to solve the problem in the next step. On the other hand, students with rubric score 1 can readily interpret the relevant information and, once accessed, immediately solve the problem.

While analyzing individual student performances helps to understand student strategy and comprehension, following the progress of strategies over time may be even more important in the learning process. For example, does a student improve his or her strategy over time and how? Does a student move from not knowing what information is relevant/not solving the problem (rubric 5), to perhaps knowing what information is relevant but not knowing how to interpret it/not solving the problem (rubric 4), and ultimately to knowing what information is relevant and how to interpret it/solving the problem immediately (rubric 1)? We are currently investigating this question.

We have observed no differences in performances among 10th, 11th, or 12th graders, and likewise among regular, honors, or AP courses. The content was novel for all the students so that large differences might not be expected. However, a tendency did exist in that students in honors and AP classes solved more problems correctly than those in regular classes. Since the current database included 470 performances, it may be that a larger database might be able to distinguish statistical differences. We will pursue this as the database grows larger. However, it is interesting to note that when AP and honors classes were compared in a related study across many problem sets, the AP students performed significantly better than the honors students ($p = 0.001$) (22).

In general, educators are becoming more familiar with the problem-based approach for instruction and are seeking such methods to use in the classroom. For those interested in authoring problem-solving scenarios to meet the needs of their curriculum, the IMMEX Laboratory offers summer and winter workshops for teams of educators representing all levels of education (for contact information see http://www.immex.ucla.edu/TopMenu/Contact.htm).

ACKNOWLEDGMENT

REFERENCES

1. **Arocha, J. F., V. L. Patel, and Y. C. Patel.** 1993. Hypothesisgeneration and the coordination of theory and evidence in novice diagnostic reasoning. Med. Decis. Making. **13:**198–211.

2. **Bransford, J. D., A. L. Brown, and R. R. Cocking (ed.).** 1999. How experts differ from novices, p. 19–38. How people learn. National Academy Press, Washington, D.C.

3. **Casillas, A. M., S. G. Clyman, G. Stephen, Y. V. Fan, and R. H. Stevens.** 2000. Exploring alternative models of complex patient management with artificial neural networks. Adv. in Health Sci. Educ. **5:**23–41.

4. **Dhillon, A. S.** 1998. Individual differences within problem-solving strategies used in physics. Sci. Educ. **82:**379–405.

5. **Fatemi, E.** 1999. Building the digital curriculum. Education Week/Technology Counts '99. **19:** 23.

6. **Lawton, M.** 2000. Making the most of assessments. Technol. Counts '98 **18:**59–62.

7. **National Research Council.** 1996. National science education standards. National Academy Press, Washington, D.C.

8. **Palacio-Cayetano, J.** 1997. Ph.D.Thesis. Problem-solving skills in high school biology: the effectiveness of the IMMEX problem-solving assessment software. University of Southern California, Los Angeles.

9. **Palacio-Cayetano, J., R. D. Allen, and R. H. Stevens.** 1999. Computer-assisted evaluation—the next generation. The Am. Biol. Teacher **61:**514–522.

10. **Palacio-Cayetano, J., S. Kanowith-Klein, and R. Stevens.** 1999. UCLA's outreach program of science education in the Los Angeles schools. Acad. Med. **74:**348–351.

11. **Refenes, A. N., N. Zapranis, and G. Francis.** 1994. Stock performance modeling using neural networks: a comparative study with regression models. Neural Networks **7:**375–388.

12. **Reggie, J.** 2000. Neural computation in medicine. Artificial Intelligence in Med. **5:**143–158.

13. **Roth, W. M.** 2000. Artificial neural networks for modeling knowing and learning in science. J. Staff Dev. **37:**63–80.

14. **Rumelhart, D. E., and J. L. McClelland.** 1986. Parallel distributed processing: explorations in the microstructure of cognition. MIT Press, Cambridge, Mass.

15. **Sivaramakrishnan, M., J. F. Arocha, and V. L. Patel.** 1998. Cognitive assessment and health education in children from two different cultures. Soc. Sci. Med. **47:**697–712.

16. **Stevens, R., J. Ikeda, A. Casillas, J. Palacio-Cayetano, and S. Clyman.** 1999. Artificial neural network-based performance assessments. Computers in Hum. Behavior **15:**295–313.

17. **Stevens, R. H.** 1991. Search path mapping: a versatile approach for visualizing problem-solving behavior. Acad. Med. **66:**S73–S75.

18. **Stevens, R. H., A. R. Kwak, and J. M. McCoy.** 1989. Evaluating preclinical medical students by using computer-based problem-solving examinations. Acad. Med. **64:**685–687.

19. **Stevens, R. H., and A. C. Lopo.** 1994. Artificial neural network comparison of expert and novice problem-solving strategies. Proc. Annu. Symp. Comput. Appl. Med. Care. **1**(Suppl.)**:**64.

20. **Stevens, R. H., A. C. Lopo, and P. Wang.** 1996. Artificial neural networks can distinguish novice and expert strategies during complex problem solving. J. Am. Med. Inform. Assoc. **3:**131–138.

21. **Stevens, R. H., J. M. McCoy, and A. R. Kwak**. 1991. Solving the problem of how medical students solve problems. MD. Comput. **8:**13–20.

22. **Stevens, R. H., T. Vendlinski, J. Palacio-Cayetano, J. Underdahl, P. Paek, M. Sprang, and E. Simpson.** 2000. Developing and implementing a K-12 technology program of case-based reasoning. Manuscript in preparation.

23. **Trotter, A.** 1998. Putting school technology to the test. Education Week/Technology Counts '98. **18.** 1.

24. **Van Melle, E., and L. Tomalty.** 2000. Using computer technology to foster learning for understanding. Microbiol. Educ. **1:**7–13.

---

*Corresponding author. Mailing address: Department of Medicine, UCLA School of Medicine, Box 957144, Los Angeles, CA 90095-7144. Phone: (310) 206-7067. Fax: (310) 267-0395. E-mail: skklein@ucla.edu.