A Tool for Helping Veterinary Students Learn Diagnostic Problem Solving

☐ Jared A. Danielson
Holly S. Bender
Eric M. Mills
Pamela J. Vermeer
Barbara B. Lockee

This study describes the result of implementing the Problem List Generator (PLG), a computer-based tool designed to help clinical pathology students learn diagnostic problem solving. Participants included 507 veterinary students: 173 in the treatment groups and 334 in the nontreatment (comparison) group. The comparison students did not use the PLG; one experimental group participated in PLG-based case-discussion sessions, and the other used the PLG both for case-discussion sessions and for homework. Both treatment groups scored significantly higher on the final exam (p = .001 and .000 respectively) than the comparison group. The findings suggest that student problem-solving ability improved because students identified all relevant data before attempting to provide a solution, and because student and expert solutions to identical problems were generated and compared using the same process.

☐ You aren't feeling well, so you go to your doctor. A physical exam reveals nothing obvious, so blood and urine samples are sent to the lab. When the results come back, your doctor has a problem. The extent to which that problem is resolved satisfactorily (and, perhaps, your health) depends on your doctor's ability to interpret all the data that have been collected. This skill, the process of turning all the information available to the physician into a diagnosis and treatment plan, is known as *clinical problem solving;* the subproblem of correctly interpreting the clinical laboratory data is *diagnostic problem solving.*

Smith and Ragan (1999) defined problem solving as "the ability to combine previously learned principles, procedures, declarative knowledge, and cognitive strategies in a unique way within a domain of content to solve previously unencountered problems" (p. 132). Jonassen (2000) suggested that problem solving involves two processes: (a) the construction of a mental model of the problem (the problem space), and (b) activity-based manipulation of the problem space. He argued that general models of problem solving have proven inadequate for dealing with the rich diversity of problems faced by learners, as manifested by the fact that, although researchers tend to value problem solving as a learning outcome, many instructional design models and teaching theories provide sparse or nonspecific guidance when it comes to instructional strategies for learning problem solving. There are many kinds of problems, varying in complexity and inextricably tied to the knowledge domains associated with them.

Clinical problem solving and processes for teaching clinical problem solving frequently have been the object of inquiry in medical education. Expert medical practitioners in the field appear to automate expert tasks into *scripts*, obscuring underlying reasoning. This has caused some researchers to suggest that rather than spending excessive time learning underlying concepts and rules, students should focus on learning many diverse clinical scenarios and solutions (Schmidt, Norman, & Boshuizen, 1990), thereby acquiring these scripts themselves directly. Other research suggests that those who are best equipped to solve clinical problems do indeed have, and access, robust structural knowledge in this domain (Bordage, 1994; Bordage & Lemieux, 1991). The latter view is commonly held among those who make a general study of teaching and learning (Gagné, Briggs, & Wager, 1992; Jonassen, 2000; Smith & Ragan, 1999).

In medical education, research suggests that the two prominent curricula, problem-based learning (PBL) and the traditional lecture-based curriculum, produce distinct problem-solving processes (Patel, Groen, & Norman, 1991). PBL generally involves presenting a group of students with a problem and requiring them to find the resources and knowledge required for solving it without formal instruction (presentation) of the underlying rules and concepts. Although the instructional presentation does not focus on component concepts, the research and problem-solving process is consumed by immersion in those concepts. Thus, PBL can be said to focus on the discovery of underlying rules and concepts in a problem-based environment. The traditional curriculum, on the other hand, tends to focus on presentation and memorization of correlations between medical conditions, causes, and treatments (Albanese & Mitchell, 1993; Berkson, 1993; Norman & Schmidt, 1992; Rivarola, Bergesse, Garcia, & Fernandez, 1997; Rosing, 1997; Wilkerson, Hafler, & Liu, 1991). The PBL approach tends to result in students who use *backward* reasoning: They use comprehension of underlying rules and concepts to test hypotheses and arrive at solutions. Lecture-based medical curricula, on the other hand, tend to result in more *forward* reasoning: When given a problem, students search their memories for a similar case, and apply it to the problem (Patel et al., 1991). Although the PBL approach might be argued to result in reasoning that is more effective for problem-solving (which inherently involves the unknown, to a certain extent), one potential disadvantage of the PBL approach is that students tend to spend more time exploring erroneous hypotheses than those who use the traditional approach (Patel et al., 1991).

Mandin, Jones, Woloschuk, and Harasym (1997) proposed a teaching strategy to benefit from the advantages of the PBL approach while side-stepping the disadvantage of straying down many errant logical paths. They suggested providing an *expert scheme* in the context of a problem when teaching diagnostics to students. This expert scheme is a framework or process for dealing with data; learners associate subordinate concepts and skills within that process or framework. Mandin et al. asserted that if students learn knowledge in the context of an expert scheme, they will be more likely to remember the knowledge as it relates to the overall problem. This expert scheme might be described as a representation of the expert problem-space–problem-solving process. This perspective is also reminiscent of Collins, Brown, and Newman's popular concept of *cognitive apprenticeship* (1989). Among other things, Collins et al. recommended providing control strategies (the problem-solving process is controlled—includes diagnosis, monitoring, and remediation), articulation (students are required to articulate their knowledge), reflection (novice and expert solutions and processes can be compared), and appropriate sequencing (problems increase in complexity and diversity).

The project described here is the result of a problem of our own: trying to help students enrolled in their second year of veterinary school to gain diagnostic problem-solving skills in the domain of clinical pathology. We set out to create a computer-based instructional tool that would provide expert guidance, and, thereby, we hoped, increase student problem-solving ability.

To avoid reviewing basic principles and

processes of instructional design that will be familiar to most readers, while at the same time providing adequate detail for intelligent interpretation of our results, we will report our development process in terms of Briggs's(1984) Culture Four criteria for research in instructional technology, as suggested by Driscoll and Dick (1999). These criteria call for researchers to (a) accurately classify the learning outcome being studied and supply objectives and test items, (b) use real curriculum materials, (c) use materials that have been systematically designed and formatively evaluated, and (d) use tests that really measure the ability to classify examples and nonexamples of the concepts. We will briefly discuss elements of the instructional design to satisfy the first three criteria. The last is presented in the Instruments section of the Methodology.

LEARNING OUTCOMES

The instructional analysis involved many hours of interviews with experts and novice learners in clinical pathology. Most of these interviews entailed giving a representative medical problem to an expert or novice, and watching and recording their process for dealing with it (while asking questions). These interviews revealed an expert process that involved, (a) identifying relevant history data, (b) identifying abnormal laboratory data, and (c) organizing data by *causal mechanism* (defined as a disruption of normal physiology). The subparts of this process are articulated as learning objectives in the hierarchical instructional analysis found in Figure 1, and categorized using Gagné's (Gagné et al., 1992) taxonomy. This set of objectives identifies what learners will do to solve a specific case, given the prerequisite domain knowledge. An instructional analysis involving all the prerequisite domain knowledge would obviously be extremely extensive.

CURRICULUM MATERIALS

In developing the strategy, we turned to the instructional analysis, producing an expert scheme (Mandin et al., 1997) that would provide what might be considered an effective environment for enhancing the development of the learner problem space. Students are provided with a framework or process for dealing with data, and associating subordinate concepts and skills with that process or framework. The computer-based instructional tool we developed is known as the Problem List Generator (PLG). The PLG interface embodies the instructional strategy. The four interaction windows that provide the essential core function of the PLG are described as follows: (a) Figure 2 contains relevant signalment (species or breed characteristics), history, and physical exam data, which together constitute the physical description of the patient and events leading up to the patient's being seen by the clinician. From the window shown in Figure 2, students identify and record relevant information, either by highlighting and pasting text directly from the paragraph narrative, or by typing their own descriptors into a dialogue box.

(b) Once students have identified the data they consider to be relevant from the window shown in Figure 2, they move to the second of the two data presentation windows, shown in Figure 3. This window presents the laboratory tests and results, and is designed to mirror a typical lab data sheet as might be seen by the student in practice. The student uses the middle pull-down menus to indicate whether or not the data fall within the normal range. If the data do not fall within the normal range, in the right-hand column, under "abnormality name," the student names the data abnormality. Most data abnormalities have a limited number of accepted names, one of which the student must enter correctly in order to move forward. The student must also identify all abnormal data before moving to the next step of the process. Prompts are given if students make three unsuccessful attempts either to name a data abnormality, or to move on to the next window, so that students do not become stuck.

(c) After students have identified and correctly named all data abnormalities, they proceed to the *construct problem list* window, as seen in Figure 4. In the right column of this window appear the observations and data abnormalities iden-

Figure 1 ☐ Hierarchical analysis of learning objectives with learning outcome classification in brackets.

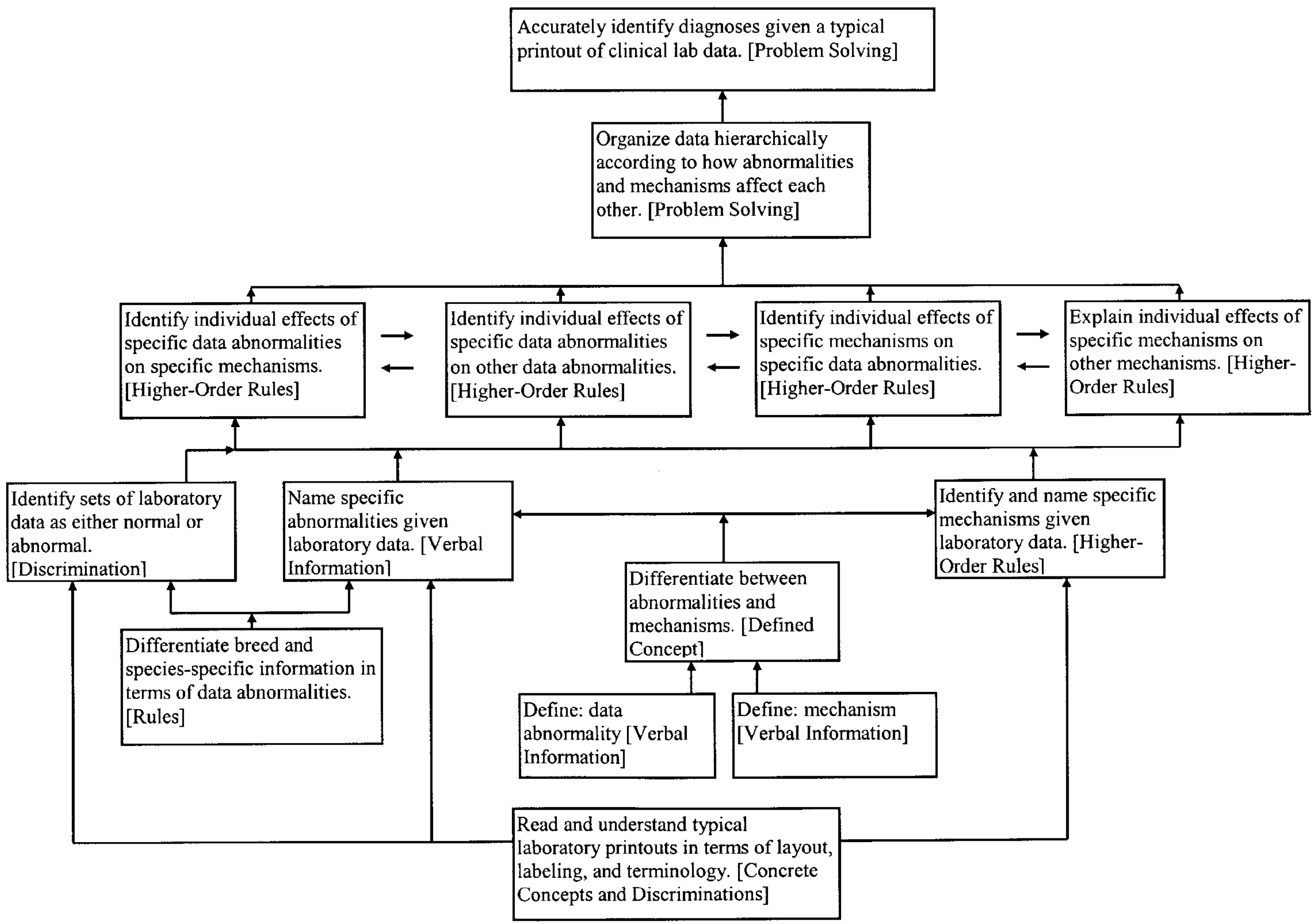


Figure 2 ☐ First data presentation window.

tified previously by the students. In the left-hand grid, learners drag and drop data elements to arrange them in a hierarchical outline format; that which is above and to the left is shown to cause (or be supported by) that which falls below and to the right. The result is called a *problem list* and represents the case solution. Here the learner identifies mechanisms of disease, as based on prerequisite knowledge of physiology and pathology. By creating the problem list, learners unambiguously communicate their understanding of relationships between data abnormalities.

(d) After completing the problem list to their satisfaction, learners make a diagnosis, and the problem list is submitted for credit. At this point, students compare their problem list to the expert problem list, as seen in Figure 5. Note that the expert problem list is identical in format to the student problem list, except that mechanisms are coded as core (new, and essential to this case), review (previously encountered essential mechanisms that are reviewed here), or framing (mechanisms that have not yet been formally covered in the course, but will be covered later, or are not central to the study of clinical pathology). The PLG contains other features and functionality, but the windows contained in

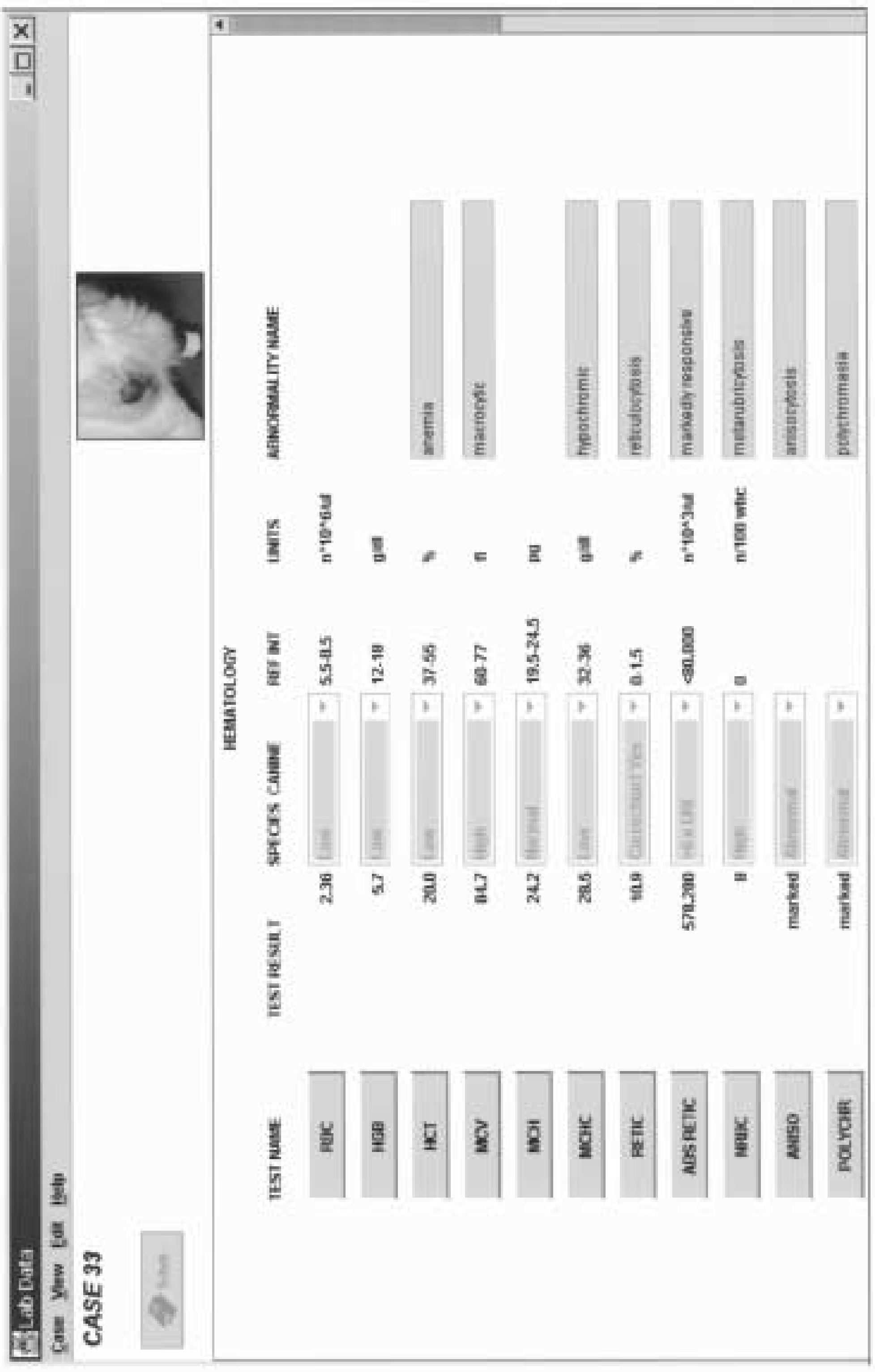
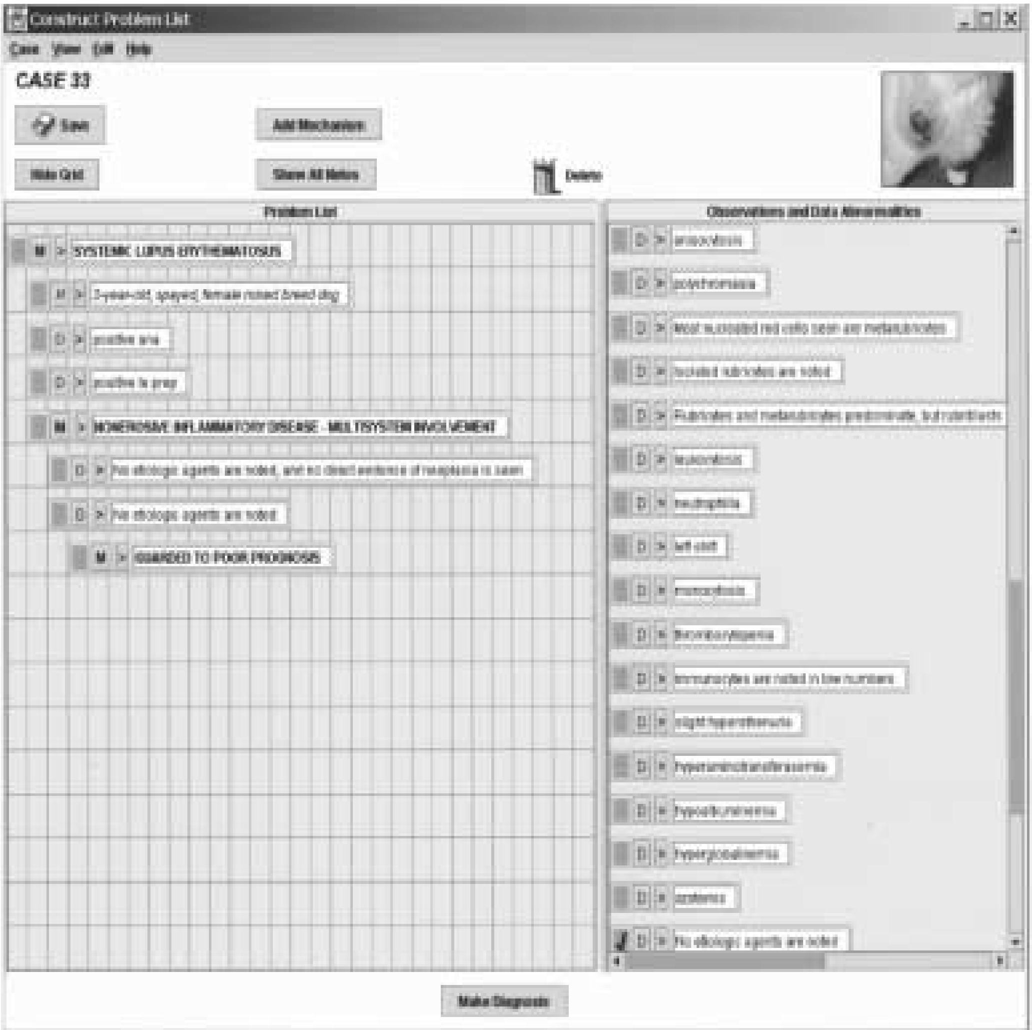


Figure 3 □ Second data presentation window—laboratory data.

Figure 4 ☐ Data synthesis.



Figures 1–5 constitute the core instructional functionality of the program.

SYSTEMATIC DESIGN, DEVELOPMENT, AND EVALUATION

Elsewhere, Danielson (1999) described the design, development, and formative evaluation of the PLG in great detail. In summary, the PLG accompanies a preexisting course with a history of success and popularity with students and alumni. The course for which the PLG was designed uses a mechanism-based approach to diagnostic reasoning, as described by Bender et al. (2000). The PLG was designed to provide a means for students to complete and receive credit for their work, and to improve their diagnostic problem-solving skills in clinical pathology. We used Dick and Carey's (1996) instructional design model to guide the design and development process of the PLG. In addition to the Dick and Carey model, we followed Tessmer and Richey's (1997) more detailed contextual analysis procedure. Because our tool was computer based, we combined the overall design process with Hix and Hartson's (1993) usability design model. A more in-depth discus-

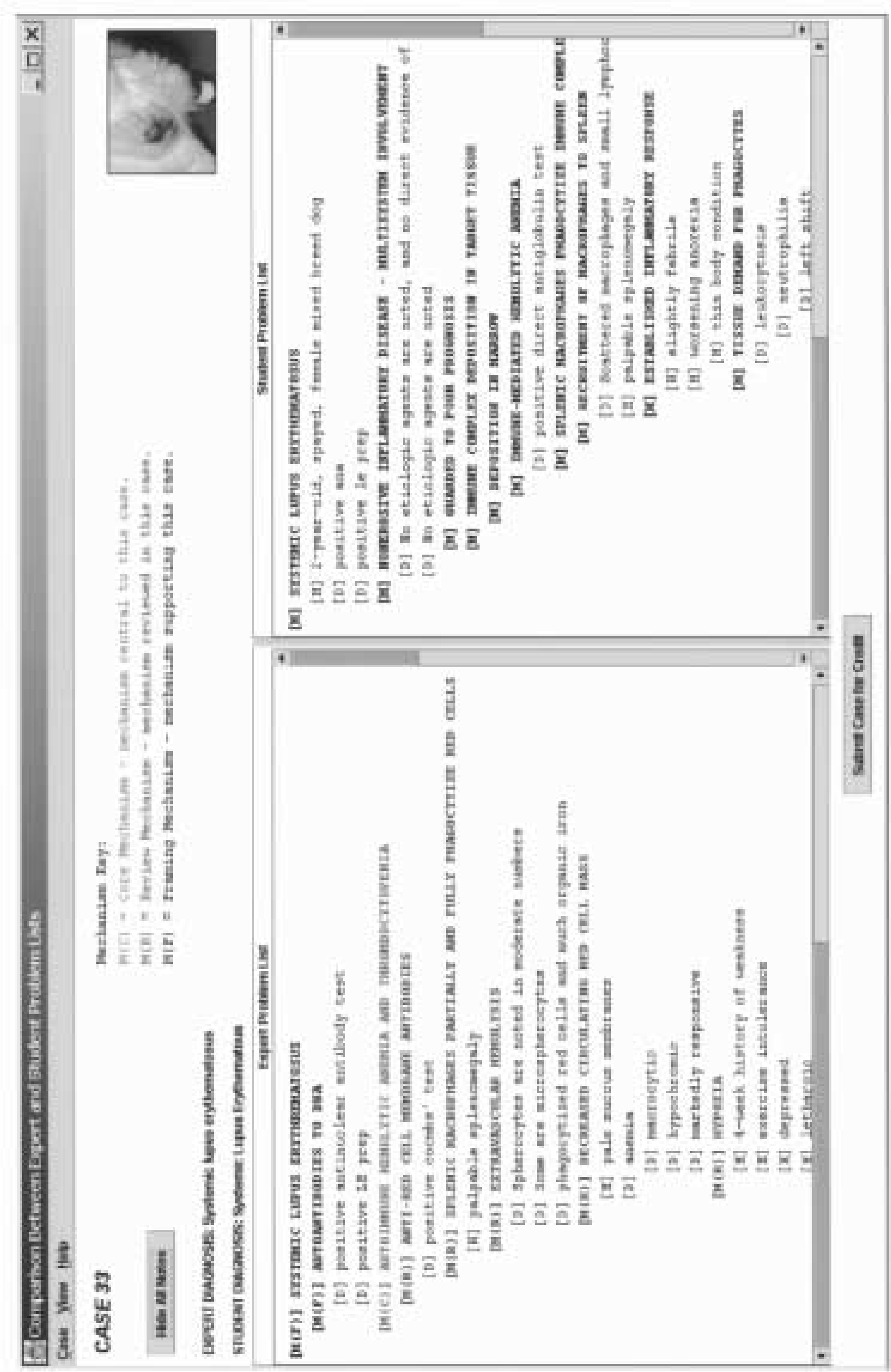


Figure 5 ☐ Feedback.

sion of how interface design can be incorporated into instructional design is provided elsewhere (Danielson, Lockee, & Burton, 2000). Finally, we bolstered the formative evaluation process with recommendations from Tessmer (1993). Here we report the results of the summative evaluation.

METHOD

Variables

The dependent variables were learning impact and usability. Usability was subdivided as clarity and feasibility. The final exam was used as one indicator of learning impact. A questionnaire given in 2001 and an interview with the instructor were also used to indicate impact (self-report). Clarity and feasibility data were gathered from the questionnaire.

Participants

There were 507 participants—all the students who participated in VM 8414 (Clinical Pathology) between 1996 and 2001. All students had an undergraduate degree, and were recruited to the veterinary college using the same entrance requirements (G.R.E. scores, GPA, and entrance interview), and drawing from the same populations. Eighty-nine percent of each class's students were recruited from the two states with which the college is affiliated. The remaining 11% were recruited from other states. Each class was approximately 70% female and 30% male, and comprised between 77 and 89 students.

Procedures

The Clinical Pathology course was taught using the following format throughout the period of the study: The course is designed around 21 lectures, 49 case discussion periods, 21 unannounced quizzes, and 90 case-based homework assignments. Lectures present component knowledge and skills necessary to understand the case homework. The unannounced quizzes are based on case homework assignments, and provide an incentive for completing the homework. The 49 case discussion periods, which are dispersed throughout the semester, are used to analyze the cases after they have been completed as homework by the students. Each student prepares one case for presentation to the class during one of the case discussion periods, and discusses that case with an instructor prior to presenting it. Instructors present lectures, lead case discussion periods, and work individually with students to prepare their assigned case discussion presentations.

During 1996–1998, the course was taught as described above. Students were required to do homework, but did not turn it in or get credit for it. The pop quizzes, which amounted to half of the total grade, provided incentive for students to complete homework case assignments.

The course was taught in 1999 just as it had been taught in 1996–1998 with three exceptions. First, during 1999, the PLG design was finalized and the formative evaluation was completed. Some students in the 1999 class were recruited to participate in formative evaluation sessions of the PLG, but exposure to the PLG was minimal, involved primarily easy or previously seen cases, and primarily occurred after the majority of the course had been taught. Students completed their case homework on paper as had been done in previous years. Second, 1999 differed from previous years in that the students were required to submit their homework assignments, and were given credit for doing so (33% of the grade was now based on homework completion, 33% on pop quizzes, and 33% on the final exam). The homework assignments were not evaluated; students were given full credit for simply turning them in. Because of the regular pop quizzes, it was not hypothesized that requiring homework to be turned in would affect student compliance with homework assignments—pop quizzes had already had that effect. Finally, prior to 1999 one of the two VM 8414 instructors retired, so a new faculty member was hired to teach one third of the lectures. One instructor, who had taught the entire time, remained and taught two thirds of the lectures. Again, it was not thought that this change would affect student outcomes because the course retained the primary instructor, and the format did not change.

In 2000 the course was taught as it had been taught in 1999, except that students were, for the first time, given the option of using the PLG to complete their homework assignments, though they could continue to complete their assignments on paper if they wished. Bugs in the PLG persisted, so only 10 of the 84 students used it to complete their homework. In 2000, however, the professor began to use the PLG to prepare and present case discussion lectures, and to prepare expert case solutions (to which students could compare their case solutions.)

In 2001, the course was taught as it had been taught in 2000, with two differences. First, the PLG was now more completely debugged, so most students used it to complete their homework. Of the 89 students in the class, 51 (59%) reported using the PLG for all of their cases. An additional 35 (21%) reported using the PLG for 90+% of their cases. Of the remainder of the class, only 9% (8 students) reported using the PLG for fewer than half of their cases. Second, because of the unexpected illness of one of the course instructors, the course was taught entirely by the first instructor (who initiated the development of the PLG, taught 50% of the lectures in 1996–1998, and taught 66% of the lectures in 1999–2000).

Instruments

The course final exam, used to measure learning impact, contains laboratory, signalment, history, and observation data for eight clinical cases. Five to seven multiple choice items accompany each case, in formats that include choosing the correct answer, choosing the incorrect answer, selecting all that apply, and so forth. An example of items from the final exam is found in Figure 6. Final exams are administered in class; they are equivalent from year to year, and are collected to prevent students from passing final exams on to subsequent classes.

The primary course instructor was interviewed informally on a number of occasions after 2000 and during 2001 to determine her impression of any effect that implementing the PLG might be having. Because these were informal conversations, no interview protocol is provided.

All 2001 students were asked to complete a survey instrument designed to assess their perception of the impact, clarity, and feasibility of the PLG. The survey instrument and results are discussed in the Results section.

RESULTS

Final Exam Data

Table 1 contains the means that were used in the year-to-year comparisons. We used a two-tailed independent samples *t* test to make all means comparisons. The significance level used for all comparisons was .05, and *p* values are provided with all comparisons. Power calculations were based on an effect size of .3. This effect size was derived from the college grading scale. Using that scale, the difference between one grade and another (e.g., *A* and *A*–) is 3 percentage points. Although a rather arbitrary grading scale is not particularly indicative of knowledge, these differences have clear significance to the students. Students are aware of the grade cutoffs, and will often study with the goal in mind of improving a grade by one or more increments; therefore, we determined that one grade increment is a justifiable target effect size. In our context, a grade increment is, by definition, 3 percentage points, which also happens to be .3 *SD* for our population of 507 students. Therefore, for differences that are found to be significant, we consider an effect size of .3 or larger to be noteworthy.

Table 1 ☐ Means and Standard Deviations for year-to-year comparisons.

Year(s)	*M±SD*	*n*
1996–1998	81.1±10.1	245
1999	82.9±8.9	89
1996–1999	81.6±9.8	334
2000	85.7±10.3	84
2001	87.3±10.1	89

Data from years prior to 2000

Because there was no reason to suspect any systematic differences between 1996, 1997, and

1998, the data from these years were combined for comparison to treatment years. Two minor changes, (one instructor change and a grading policy change, as described in the Procedures section), were introduced in 1999. We did not think these changes would prove to be significant. The final exam score difference between the 1996–1998 group and the 1999 group did not, in fact, prove to be statistically significant, $t = 1.47$, $p = .142$, power = .67. Data from 1999 were, therefore, clumped with the 1996–1998 data for comparison to the years in which the PLG was implemented.

Data from 2000

The difference between final exam scores for 2000 ($M = 85.7$) and 1996–1999 ($M = 81.6$) was statistically significant, $t = 3.47$, $p = .001$, effect size (Cohen's d) = .42.

Data from 2001

The difference between final exam scores for 2001 ($M = 87.3$) and 1996–1999 ($M = 81.6$) was statistically significant, $t = 4.94$, $p = .000$, effect size (Cohen's d) = .58. The difference between exam scores for 2000 and 2001 was not statistically significant, $t = 1.04$, $p = .302$, power = .50.

Faculty Member Interview

The faculty member teaching clinical pathology, and who initiated the development of the PLG, was interviewed to determine her impressions of the effect of the PLG, if any, on learning. Three main points emerged. First, she indicated that student problem lists and reasoning appear to be better since implementation of the PLG. Students seem to struggle less and be less frustrated. Meetings with students prior to their case presentations have become much shorter and students now come prepared with more defensible and logical problem lists. Second, she indicated that her own problem lists (the expert list used for case discussions and PLG comparison) have become more precise and consistent. Finally, she indicated that her problem lists have become more detailed and complete. She did not indicate any disadvantages to PLG use.

Survey Data

The quantitative results for items 1–16 of the 2001 survey data can be found in Table 2. These items seek to determine student perceptions of PLG use in terms of clarity, feasibility, and impact, and use a Likert-style scale of 0 to 10. An additional 5 Likert-style items, asking for general impressions of the PLG, and providing students with antonymic descriptors (i.e., *wonderful* = 10 vs. *terrible* = 0; and *interesting* = 10 vs. *uninteresting* = 0) were also included. These 5 items compared favorably with the item means found in the other 16 items. Means for these items favored the PLG, the lowest being 8.8 (SD = 1.2) and the highest being 9.4 (SD = 1). In addition to the Likert items, the survey also included 5 free-text items. Responses to those items are summarized as follows:

Item 22 asked that students indicate why they ranked specific items particularly negatively. There were 21 participants responding, with 3 of those indicating that they had made no negative comments. Of the remaining comments, 9 reported general technical problems, 8 reported frustration that they had to be connected to a network to run the PLG, 2 indicated that too many cases were required, 1 commented that "it makes you learn, which is hard," and 1 commented that it was difficult to learn to use.

Item 23 asked students to indicate what they liked best about using the PLG. The 72 participants responding made 152 comments in all. Of those, 36 responses indicated that the PLG helps organize thoughts or data, 26 indicated that the PLG enhances completeness, 25 favored the expert feedback, 18 indicated that the PLG enhances memorization of data abnormality names, 17 indicated that the PLG is easy to use, 13 indicated that the PLG is convenient (e.g., that they could complete and turn in homework electronically), 8 were general statements about the benefit for learning (the PLG makes things "make sense"), 5 favored the pictures, 3 favored the fact that cases are used, and 1 indicated that the PLG helps class presentation.

Item 24 asked students to indicate what they liked least about using the PLG. The 72 responding participants made 82 comments in all. Of

Figure 6 ☐ Final exam excerpt.

You are presented with a 5-year-old, spayed female poodle. The owner complains that the dog just "isn't right." He says that he can't put his finger on when it started, but for the last 2 months or so, the dog is not eating as enthusiastically, occasionally vomits, and periodically has loose stools. Two weeks ago, she seemed brighter, but now she is progressively weaker, lethargic and occasionally urinates on the living room rug. On physical examination, you note a depressed, dehydrated patient, with a weak femoral pulse, bradycardia, and hyperpnea.

		HEMATOLOGY	Canine Ref. Interval	Units
RBC		6.10	5.5-8.5	n x 10 ^{6}/µl
HGB		14.1	12-18	g/dl
HCT		42.0	37-55	%
MCV		68.9	60-77	fl
MCH		23.1	19.5-24.5	pg
MCHC		33.6	32-36	g/dl
RETIC			0-1.5	%
ABS RETIC			<80.000	n x 10 ^{3}/µl
NRBC			0	n/100 wbc
ANISO				
POLYCHR				
HYPOCHR				
POIK				
COMMENT				
WBC		10.100	6.000-17.000	n x 10 ^{3}/µl
SEG	49%	4.949	3.000-11.400	n x 10 ^{3}/µl
BAND	0%	0.000	0-0.300	n x 10 ^{3}/µl
LYMPH	43%	4.343	1.000-4.800	n x 10 ^{3}/µl
MONO	2%	0.202	0.150-1.350	n x 10 ^{3}/µl
EOS	6%	0.606	0.100-0.750	n x 10 ^{3}/µl
BASO		0.000	rare	n x 10 ^{3}/µl
META		0.000		n x 10 ^{3}/µl
MYEL		0.000		n x 10 ^{3}/µl
OTHER		0.000		n x 10 ^{3}/µl
PLATELETS		440	200-900	n x 10 ^{3}/µl
PLT EST		adequate		
PP		8.8	6.0-7.5	g/dl
WBC MORPH				

URINALYSIS (Cystocentesis)

		SEDIMENT	
COLOR	straw		
TRANSPARENCY	clear	RBC/HPF	neg
SPECIFIC GRAVITY	1.021	WBC/HPF	0-2
pH	6.0	CASTS/LPF	neg
PROTEIN	neg		
GLUCOSE	neg	EPI CELLS/HPF	occasional transitional
KETONES	neg	CRYSTALS	amorphous
BILIRUBIN	neg		
UROBILINOGEN EU/dl	normal	BACTERIA	neg
OCCULT BLOOD	neg	MISC	

Figure continues

these, 23 comments cited technical problems that were owed to the student's home or study environment (they did not have a computer, or their computer was old and slow, etc.), 22 comments had to do with technical or situational problems beyond the student's power to control (e.g., the PLG tied up phone lines, the PLG was not Macintosh compatible, etc.), 18 comments related to interface or functionality problems (e.g., students could not select multiple lines in the problem list constructor window), 12 comments had to do with factors that increased the difficulty of the learning task and were deliberately designed into the PLG (e.g., the PLG

Figure 6 ☐ Final exam excerpt (*continued*).

CHEMISTRY ☐ SERUM ☒ PLASMA

Specimen Comments

Test name	Results	Ref interval	Units
TOTAL PROTEIN	8.9	5.3-7.8	g/dl
ALBUMIN	4.4	2.3-4.3	g/dl
GLOBULIN	4.5	2.7-4.4	g/dl
UREA NITROGEN	79	5-28	mg/dl
CREATININE	3.1	<1.5	mg/dl
TOTAL BILIRUBIN	0.5	0.1-0.6	mg/dl
ALANINE AMINOTRANSFERASE	83	4-66	mU/ml
PHOSPHORUS	5.5	2.5-5.0	mg/dl
GLUCOSE	70	71-115	mg/dl
ALKALINE PHOSPHATASE	33	<88	mU/ml
SODIUM	130	145-155	mEq/l
CHLORIDE	97	112-124	mEq/l
POTASSIUM	6.1	2.7-5.0	mEq/l
TOTAL CO2	10	13-29	mEq/l
ANION GAP	29.1	10-25	mEq/l
BLOOD GASSES (arterial)			
negative log of hydrogen ion concentration	7.21	7.36-7.44	
Partial pressure of carbon dioxide	29	38-42	mmHg
BASE EXCESS	-15	-2.0 - 2.0	mEq/l
BICARBONATE	12.5	17-24	mEq/l
ENDOCRINE TESTING			
RESTING CORTISOL	0.1	0.5-4.0	µg/dl
POST-ACTH CORTISOL	0.1	8-20	µg/dl

1. Which mechanism best explains the hyponatremia in this case?
   Choose the correct answer for 1 point.
   a. The dog is sequestering sodium excessively in the peritoneal cavity.
   b. Diabetes mellitus is causing shift of intracellular water and dilution of extracellular sodium.
   c. The dog is losing sodium excessively in the urine.
   d. An acidemia is causing extracellular sodium to shift into the intracellular compartment.
2. Choose the most likely cause for the plasma glucose alteration for 1 point.
   a. The hypoglycemia is due to chronic liver failure.
   b. The hypoglycemia is due to hypocortisolism.
   c. The hypoglycemia is due to insulinoma.
   d. The hypoglycemia is due to hypoaldosteronism.
3. Choose the correct reason for the hyperkalemia for 1 point.
   a. The hypochloridemia is causing a false hyperkalemia due to an analytical error.
   b. The animal is unable to excrete potassium into the distal nephron urine appropriately.
   c. Anorexia is causing the animal to retain extracellular potassium.
   d. The diarrhea is causing extracellular potassium to shift intracellularly.
4. Choose the incorrect answer regarding the acid-base abnormality for 1 point.
   a. The dog most likely has a titration metabolic acidosis due to excessive lactate production.
   b. The dog has a titration metabolic acidosis as indicated by the increased anion gap.
   c. The data show a respiratory acidosis that is most likely a primary acid base disturbance.
   d. The data show a decreased TCO2 which is an estimate of decreased plasma bicarbonate.
5. Choose the incorrect explanation for the urine specific gravity for 1 point.
   a. The urine specific gravity is a normal finding (consistent with health) in this patient.
   b. The urine specific gravity implies renal medullary washout (decreased renal medullary osmolality) in this patient.
   c. The animal has a decreased capacity to reabsorb sodium from the renal tubular filtrate.
   d. Your patient's kidneys have a decreased capacity to concentrate urine.

Table 2 ☐ Means, standard deviation, median, maximum, and minimum by item for 2001.

Questionnaire item:		*M*±*SD*	*Mdn*	*Max.*	*Min.*	*n*
1. Indicate the approximate percentage of cases on which you used the PLG. Include cases you worked on in a group, where the group used the PLG.	0% 100% 0 1 2 3 4 5 6 7 8 9 10 NA	8.7±2.5	10	10	1	86
2. When I use the Problem List Generator, I feel . . .	Overwhelmed Comfortable 0 1 2 3 4 5 6 7 8 9 10 NA	9.1±1.2	9	10	4	80
3. Learning how to use the PLG was . . .	Difficult Easy 0 1 2 3 4 5 6 7 8 9 10 NA	8.5±1.9	9	10	1	80
4. Navigating through a case using the PLG is . . .	Confusing Clear 0 1 2 3 4 5 6 7 8 9 10 NA	9.1±1.2	9	10	5	80
5. Using the Problem List Generator made me account for more lab data than I otherwise would have accounted for.	Less Same 0 1 2 3 4 5 6 7 8 9 10 NA	8.9±1.5	10	10	5	80
6. Using the Problem List Generator made my problem lists more precise than they would have been otherwise.	Less Same 0 1 2 3 4 5 6 7 8 9 10 NA	9.0±1.3	10	10	5	80
7. It is hard to remember what all the buttons and menus of the PLG do.	Yes / Too hard No 0 1 2 3 4 5 6 7 8 9 10 NA	9.5±1.1	10	10	5	79
8. Overall navigation (getting around in the program) is . . .	Irritating Easy 0 1 2 3 4 5 6 7 8 9 10 NA	8.9±1.5	9	10	2	80
9. I am frustrated by technical problems with the PLG.	Frequently Never 0 1 2 3 4 5 6 7 8 9 10 NA	7.7±2.0	8	10	2	80
10. The PLG makes doing my Clinical Pathology homework more enjoyable than doing it on paper.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	8.9±1.8	10	10	2	80
11. The PLG makes doing my Clinical Pathology homework more worthwhile than doing it on paper.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	8.9±1.5	9	10	3	80
12. I like being able to do my problem lists on a computer.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	9.3±1.4	10	10	1	80
13. I like being able to do my problem lists online.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	8.3±2.2	9	10	1	79
14. I like having my problem lists turned in automatically as soon as I finish them.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	9.6±1.1	10	10	5	80
15. Using the problem list generator helps me to organize my thoughts about a case.	Definitely not Absolutely 0 1 2 3 4 5 6 7 8 9 10 NA	9.1±1.4	10	10	3	80
16. Using the problem list generator makes understanding clinical pathology . . .	Harder Easier 0 1 2 3 4 5 6 7 8 9 10 NA	8.9±1.4	9	10	4	80

Note: Copyright 2002 by Jared A. Danielson, Holly S. Bender, Pamela J. Vermeer, Eric M. Mills requires correct spelling of data abnormality names), 4 comments were that too much homework is required; 2 respondents indicated that there was nothing they did not like about the PLG.

Item 25 asked participants what they would change about the PLG if they could. The 60 students who responded made 64 suggestions in all. Of these, 32 suggestions involved interface and functionality enhancements, such as making it easier to print, or to copy and paste throughout the program, 9 comments indicated that the PLG should be made "faster," 9 recommendations were that the PLG not be changed at all, 6 comments indicated that the PLG is too much work (either that it should not require correct spelling, etc., or that fewer cases should be required). There were 4 recommendations that the PLG be available on more computers at the veterinary college, 3 recommendations that the PLG be used in more classes at the college, 2 indications that respondents' computers simply did not have enough RAM to run the PLG, 1 recommendation that the PLG be expanded to include content-specific tutorials, and 1 recommendation that the PLG be made to work with Macintosh computers.

Item 26 asked for any other comments. The 32 responding participants made a total of 35 responses. Of these, 30 were either general positive statements such as "great program," or specific indications that the PLG had helped learning; 4 comments had to do with the PLG taking too much time or requiring too many cases and 1 person indicated that the PLG had lost 3 homework assignments.

Item 27 asked that those using the PLG for fewer than 20% of their cases indicate why. There were 7 respondents, and 9 responses. Of these, 5 had to do with computer access at the student's home (no computer or a slow computer), 3 were that the cases could be done faster on paper, and 1 indicated a general reluctance to use computers.

DISCUSSION

The study investigated the effectiveness of the PLG—a software tool designed to improve student problem solving in the complex domain of clinical pathology. The results of the study suggest that students learning through classroom use of the PLG, as well as students using the PLG for both classroom and homework use, earn higher final exam scores than those who do not use the PLG. It further suggests that students and instructor think that the PLG is beneficial for learning, and that the students, in general, find the PLG to be usable, both in terms of clarity and feasibility.

Learning Impact

The data suggest that those who took VM 8414 in 2000 and 2001 performed significantly better on the final exam than their counterparts in previous years. Furthermore, the effect sizes of .42 and .58 respectively appear to be noteworthy. (We provided rationale in the Method section for considering an effect size of .3 or larger to be of consequence for this study.)

The fact that the research design was not experimental increases the possibility that differences between groups other than PLG use caused the differences in final exam scores. However, all students in the study were selected to the college the same way, and group assignment was a function of nothing other than the year that the students applied for veterinary school. Therefore, because we specifically designed the PLG to have a beneficial learning impact, and because the improvement corresponds with the implementation of the PLG, we are inclined to cautiously conclude that the PLG improved learning as indicated by final exam scores. This conclusion is supported by the questionnaire and interview data. Students generally felt that using the PLG helped them to understand clinical pathology better; the instructor thought that PLG use improved her own problem lists, as well as student diagnostic problem-solving ability.

Richey and Nelson (1996) observed that with developmental research it is more likely that unanticipated events will affect research procedures than with other kinds of research. This was consistent with our experience. In this case, bugs in the software in 2000 led us to have a year of partial PLG implementation, which we had not anticipated. Rather than ignore data from 2000, we have reported the data as a nontypical implementation. In 2000, the PLG was used for case discussion in class, but only about 10% of the students used it regularly for case homework. Because we anticipated that the primary benefit of the PLG would be in its use as a homework practice tool, we would have expected exam scores for 2000 to be more similar to prior years than to 2001 (when the PLG was used in class by all students, and for homework by almost all students). The opposite was the case, however. Although 2001 scores were higher than 2000 scores, and 2000 scores were higher than previous years, the statistically significant difference was found between 2000 and previous years rather than between 2000 and 2001. This could be due to several factors. First, the benefit from the instructors using the PLG for in-class discussion could be greater than anticipated. This hypothesis would be consistent with the primary instructor's observation that using the PLG increased her detail and consistency. It seems reasonable that as detail and consistency of in-class presentations improved, student comprehension improved as well. It is also possible that a Type II error occurred, and that 2001 students really did do significantly better than their 2000 counterparts. Because the power of that particular comparison was .5, it seems possible that our research design was inadequate to detect a difference, and that we would have found a difference had there been more participants.

Clarity

The survey data suggest that, for the majority of the students, the PLG was easy to use, and easy to learn to use. This assertion is supported by responses to items 22–27, though respondents clearly felt there was room for improvement in the interface. Many of the interface improvement recommendations either have been implemented, or are in the process of being implemented. However, we have not chosen to make the PLG easier by removing learning requirements (such as the requirement for correct spelling) that were embodied in the PLG by design.

Feasibility

The fact that all but 9% of the class (eight students) used the PLG for more than half of their 90 cases in 2001 indicates that implementing the PLG was feasible. Furthermore, many students specifically commented that they felt the PLG was more convenient to use than simply doing the problem lists on paper. Finally, most suggestions for improving aspects of feasibility involved expanding or enhancing the use of the PLG (for example, making it available on more computers, or less dependent on a dedicated network line), but there were no suggestions that PLG use be discontinued.

The PLG and Current Theoretical Perspectives on Teaching and Learning

In designing the software, we relied on relevant learning theory as well as personal experience. The design process was planted solidly in the instructional systems tradition, but was, nevertheless, informed by a number of theoretical perspectives (Danielson, 1999). Of key importance in approaching the problem of teaching problem solving was our belief that problem solving is best accomplished through a thorough knowledge of the underlying rules and concepts of the domain. In clinical pathology, this philosophy is reflected in what we have called elsewhere a *mechanism-based* approach to practice and teaching (Bender et al., 2000). This mechanism-based approach is probably best described as a manifestation of backward reasoning, as discussed earlier in this paper. It would be hasty, however, to interpret these findings as exclusive support for a particular philosophical approach to teaching and learning at the expense of others. We wish to avoid the error of hurriedly planting a particular philosophical flag on the turf of a practice that might be characterized with equal persuasiveness from a number of perspectives.

For instance, in practice the PLG uses some instructional principles that are characteristic of constructivist theory. In some ways, it presents what Duffy and Cunningham (1996) referred to as "the problem as a stimulus for authentic activity" (p. 190). In this model, instead of *teaching*, the teacher-facilitator supports the student's learning "as skills are developed through working on the problem" (p. 191). At the same time, Duffy and Cunningham argued that discovery learning, scaffolding, cognitive apprenticeship, coaching, and collaborative learning, all frequently referred to as constructivist methods, are only truly constructivist to the extent that they are used not to teach students what they "should do/know and when they should do/know it" (p. 191), but rather to "support the students in developing their critical thinking skills, self-directed learning skills, and content knowledge in relation to the problem" (p. 191). In this study the PLG was used in the context of didactic content presentation (case discussion and lecture), and not in a PBL environment. Furthermore, the PLG requires that students identify all pieces of the problem prior to formally organizing the data, and that students communicate their understanding of relationships between data in a standard format. Therefore, while we have employed strategies commonly identified as constructivist, our approach is not constructivist by Duffy and Cunningham's definition. Rather than commit our findings to a particular philosophical approach, then, we will discuss how what we have found might inform design of computer-based learning tools, and then tie those principles back to broader theoretical ideas.

Implications for the Design of Computer-Based Learning Tools

What do these findings imply about how to design effective computer-based learning tools? We suggest two principles that might be generalized to the design of other computer-based learning tools.

First, in our domain, it seems to have been effective to establish a learning framework in which the subcomponent information inherent in the problem—the information that will be synthesized in arriving at the solution—is identified, in large part, prior to the synthesis–problem-solving activity. The PLG accomplishes this by requiring that all abnormal laboratory data be identified prior to the creation of the problem list, and that the problem list be completed prior to the declaration of the solution (diagnosis). Of course with complex problems in very ill-structured domains, figuring out *how to identify* relevant information might be the central problem-solving process (in which case this recommendation might not fit at all.) In this case, we worked to identify the process employed by experts in the field when encountering unique problems, and to instill that process into the software. The fact that this appears to have been successful supports the general idea that as Mandin et al. (1997) suggested, embodying expert processes or help in the form of an expert scheme can be useful. It also suggests that the guiding strategies mentioned in the initial section of the paper that is associated with concepts of cognitive apprenticeship (Collins et al., 1989) can be fruitful when implemented in a computer-based environment.

Second, students and faculty using the PLG represent their understandings of each problem in a standard format. This produces a powerful feedback mechanism. Students can see how an expert dealt with precisely the same problem, involving the same data. The expert solution is presented in the same format as the student solution. This goes somewhat beyond the well-established principle that immediate and detailed feedback promotes learning and retention. In this case, both the expert and the novice approach an identical problem, produce solutions independently, and then compare the results in an identical format. No two solutions are identical, but it is easy to identify where relationships and concepts match and where they do not. This principle is likely to have application in similar domains. Business students, for example, might be given a series of cases containing information for a company's resources and liabilities, and be required to produce a business plan (or portion thereof) for that company. On completion of the assignment, rather than receiving traditional feedback (or in addition to it), students could view a plan created by the instructor for dealing with the identical scenario. The rules for articulating the business plan would have to be fairly well established, so that differences in plans could be attributed to differences in the way the problem was approached, and not to ambiguities in communication. This general idea seems consistent with Jonassen's (2000) suggestion that "the key to learning to solve problems is the problem space construction, because rich problem representations most clearly distinguish experts from novices and scaffold working memory" (p. 82). The fact that the PLG appears to have been successful both as a presentation tool and as a practice-presentation tool suggests that it might help clarify and standardize the representation of the problem space, both for teacher and learner. Perhaps more important than the specific layout of information in the PLG is the fact that all the participants in the learning activity (teachers and learners alike) use the same set of data in constructing their explanation of the problem, and the same format for communicating it. We would hypothesize that this assists novice learners as they develop their problem space in the domain.

Limitations and Future Research

The similarities between the problems encountered through the PLG and problems encountered by practitioners are substantial. At the same time, the PLG case problems have been rendered simpler and more focused to teach clinical pathology (interpretation of laboratory data). Practitioners in the field have more data, and more decisions to deal with than the PLG provides. First, they have different (one would hope, richer) physical exam data because they personally examine the patient. Second, they have some flexibility in choosing what lab tests to run, and, therefore, might have less, more, or different lab data than what the PLG presents. Third, they may choose to gather more data from radiographs and other types of imaging, electrocardiograms, and so forth. Fourth, in many cases, they can follow a patient's progress over time. Although some of the cases in the PLG follow specific patients over a period of time, most cases involve only the one-time data that are presented. Therefore, although PLG practice transfers well to the final exam, we have not yet had the opportunity to explore how it transfers to a clinical environment.

Future research will focus on how learning problem-solving subtasks (such as clinical pathology) transfer to the macroproblem (clinical problem solving) of which the subproblems are a part, or, in other words, how well this kind of practice will translate to the broader clinical problem-solving environment. □

Jared A. Danielson [jadaniel@iastate.edu] is in the Department of Veterinary Pathology, College of Veterinary Medicine, Iowa State University, Holly S. Bender, Eric M. Mills, and Pamela J. Vermeer are also now at Iowa State University. Barbara B. Lockee is at Virginia Tech.

All of the authors were at Virginia Tech during the development of the Problem List Generator (PLG), which was supported by the USDA Higher Education Challenge program and by Virginia Tech seed funding. The contents of this article were also partially developed under a grant from the Learning Anytime Anywhere Partnerships (LAAP), a program of the Fund for the Improvement of Postsecondary Education (FIPSE), U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government. The authors wish to thank John Burton for the key role he played in the design and development of the PLG, and Jimmie Fortune for his insightful review of early versions of this manuscript.

Correspondence concerning this article should be addressed to Jared Danielson, Department of Veterinary Pathology, College of Veterinary Medicine, Iowa State University, Ames, Iowa 50011-1250.

REFERENCES

Albanese, M.A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues [published erratum appears in Acad Med 1993 Aug;68(8):615] [see comments]. *Acad Med, 68*(1), 52–81.

Bender, H.S., Lockee, B., Danielson, J., Mills, E., Boon, G., Burton, J., et al. (2000). Mechanism-based diagnostic reasoning: Thoughts on teaching introductory clinical pathology. *Veterinary Clinical Pathology, 29*(3), 77–83.

Berkson, L. (1993). Problem-based learning: Have the expectations been met? *Acad Med, 68*(10 Suppl), S79–88.

Bordage, G. (1994). Elaborated knowledge: A key to successful diagnostic thinking. *Acad Med, 69*(11), 883–885.

Bordage, G., & Lemieux, M. (1991). Semantic structures and diagnostic thinking of experts and novices. *Acad Med, 66*(9 Suppl), S70–72.

Briggs, L.J. (1984). Trying to straddle four research cultures. *Educational Technology, 24*(8), 33–34.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Danielson, J.A. (1999). *The design, development and evaluation of a web-based tool for helping veterinary students learn how to classify clinical laboratory data.* Unpublished doctoral dissertation, Virginia Tech, Blacksburg.

Danielson, J.A., Lockee, B., & Burton, J. (2000). ID and HCI: A marriage of necessity. In B. Abbey (Ed.), *Instructional and cognitive impacts of Web-based education* (pp. 118–128). Hershey USA: Idea Group Publishing.

Dick, W., & Carey, L. (1996). *The systematic design of instruction* (4th ed.). New York, NY: HarperCollins College Publishers.

Driscoll, M.P., & Dick, W. (1999). New research paradigms in instructional technology: An inquiry. *Educational Technology Research and Development, 47*(2), 7–18.

Duffy, T.M., & Cunningham, D.J. (1996). Constructivism: Implications for the design and delivery of instruction. In D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 112–142). New York: Simon and Schuster Macmillan.

Gagné, R.M., Briggs, L.J., & Wager, W.W. (1992). *Principles of instructional design* (4th ed.). Fort Worth: Harcourt Brace Jovanovich College Publishers.

Hix, D., & Hartson, H.R. (1993). *Developing user interfaces: Ensuring usability through product and process.* New York: John Wiley & Sons, Inc.

Jonassen, D.H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63–85.

Mandin, H., Jones, A., Woloschuk, W., & Harasym, P. (1997). Helping students learn to think like experts when solving clinical problems. *Acad Med, 72*(3), 173–179.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Acad Med, 67*(9), 557–565.

Patel, V.L., Groen, G.J., & Norman, G.R. (1991). Effects of conventional and problem-based medical curricula on problem solving. *Acad Med, 66*(7), 380–389.

Richey, R.C., & Nelson, W.A. (1996). Developmental research. In D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1213–1245). London: Simon & Schuster Macmillan.

Rivarola, V.A., Bergesse, J.R., Garcia, M.B., & Fernandez, A.C. (1997). A problem-based learning approach to a biological chemistry laboratory class for students of veterinary medicine. *Biochemical Education, 25*(1), 22–23.

Rosing, J. (1997). Teaching biochemistry at a medical faculty with a problem-based learning system. *Biochemical Education, 25*(2), 71–74.

Schmidt, H.G., Norman, G.R., & Boshuizen, H.P. (1990). A cognitive perspective on medical expertise: Theory and implication [published erratum appears in Acad Med 1992 Apr;67(4):287]. *Acad Med, 65*(10), 611–621.

Smith, P.L., & Ragan, T.J. (1999). *Instructional design* (2nd ed.). Upper Saddle River, New Jersey: Merrill.

Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training.* London: Kogan Page.

Tessmer, M., & Richey, R.C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development, 45*(2), 85–115.

Wilkerson, L., Hafler, J.P., & Liu, P. (1991). A case study of student-directed discussion in four problem-based tutorial groups. *Acad Med, 66*(9 Suppl), S79–81.