Helping Learners Gain Diagnostic Problem Solving Skills: Specific Aspects of the Diagnostic Pathfinder Software Tied to Learning Outcomes

Jared Danielson
Holly Bender
Eric Mills
Pamela Vermeer
Vanessa Preast
Iowa State University

Introduction

Problem solving is of critical importance in many disciplines. In medicine, the clinician's ability to arrive at the correct diagnosis often means the difference between life and death. Despite its importance and a significant amount of research regarding how to improve problem solving, few unambiguous answers have emerged for promoting problem solving learning.

This paper follows up on recent published findings (Danielson, Bender, Mills, Vermeer, & Lockee, 2003) concerning the Diagnostic Pathfinder (DP), a software learning tool designed to help clinical pathology students improve their ability to solve diagnostic problems. That paper described the functionality of the DP, and showed that using the DP for homework and lecture improved students' grades on a case-based final exam by a full letter grade. Very little was discussed in that paper as to how characteristics of the software supported the gains that were seen. In the two years since that report was written, the DP has been used by 640 more students at five college of veterinary medicine. Quantitative learning gains similar to those reported in 2003 have been observed (as yet unpublished) in two other settings. In this paper we turn our attention to explaining those gains, both theoretically, and in terms of the qualitative data we have been able to collect from learners over the past two years. Our goal is to associate those gains with specific characteristics of the DP by and meaningfully categorizing and characterizing thousands of comments from students who used the DP in a number of different settings, and to use those comments to illustrate, from the students' perspective, how using the DP accomplished what it did. These ideas will be tied to current theory regarding the teaching and learning of problem solving.

Defining Problem Solving

One difficulty in building upon current problem solving research is that researchers in various fields study problem solving for differing purposes and define problem solving in different ways. Even within individual domains, the conceptual waters are muddy. In the domain of cognitive psychology, for example, where problem solving has received considerable attention, the heart of the issue seems to lie in whether to define problem solving in terms of the "gap" between what the problem solver knows and what he/she must figure out to solve a problem (Wenke & Frensch, 2003), or in terms of measurable characteristics of the problem solving task (Quesada, Kintsch, & Gomez, in press). Each class of definitions is both useful and problematic. Defining problem solving purely in terms of addressing the gap between what the learner knows and what he/she needs to know implies that literally any task can constitute problem solving (or not), depending on the interaction between the problem solver and the presented problem. This state of affairs makes it difficult to provide "controlled," yet authentic, instances of problem solving for laboratory study. At the same time, the skill that problem solvers employ when they address gaps in their own knowledge seems a worthy object of study, and it seems short-sighted to ignore this phenomenon simply because it is difficult to control for in a laboratory setting. As reiterated by Ericsson (2003), it has long been recognized that one potential pitfall of experimental science is that simple and convenient forms of inquiry can prove inadequate for studying complex phenomena.

If the definition of problem solving is unclear in the field of psychology, it is more so across many of the disciplines that attempt to teach students how to solve problems. While some researchers in various teaching fields (such as biology (M. U. Smith, 1991a) and chemistry (Bodner, 1991)) have made laudable efforts to put forward unified definitions of problem solving, one disturbingly prevalent practice seems to be to discuss problem solving in terms of whatever experts do in any given field when confronted with a problem that nonexperts can't solve. Researchers who do this face the problems embodied in both of the definitions put forward by cognitive psychologists without reaping any of the benefits of either; problem solving thereby ends up being neither defined by task characteristics nor by problem solver characteristics, and hence literally can mean almost anything to anybody. (This problem can be observed by perusing the chapters of Smith's (Ed.) *Toward a Unified Theory of Problem Solving: Views From the Content Domains* (M. U. Smith, 1991b).)

Researchers in the domain of instructional design generally have committed to a gap-based definition of problem solving (e.g. (Gagné, Briggs, & Wager, 1992; Jonassen, 2000; P. L. Smith & Ragan, 1999)). Task-specific distinctions are considered (perhaps best illustrated in Jonassen's proposed problem-solving typology (Jonassen, 2003)), but the a fundamental characteristic of addressing the unknown in some way remains a key component of problem solving. Smith and Ragan's (1999) definition of 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), is typical. For the purpose of our discussion, we will align ourselves with this definition because it most closely characterizes the phenomenon we wish to discuss. Specifically, for the purposes of this proposal, *problem solving* occurs when a person employs existing knowledge and skills to achieve a goal state that he or she never before had achieved. If successful, the problem solving process results in the learning of new knowledge/skills relevant to solving the problem that was encountered, and enables the problem solver to deal with similar future situations without having to go through the problem solving process again. How, then, is an instructional designer to improve individuals' abilities to do what they don't know how to do? That is the challenge we address here. To provide a framework, we will first discuss diagnostic problem solving from the perspective of medicine.

Medical Diagnostic Problem Solving

Medical problem solving literature lacks consensus regarding the processes and knowledge structures that contribute to diagnostic problem solving. This is partly due to the fact that, as discussed earlier, definitions of the term "problem solving" vary and the term generally tends to be used synonymously with "expert performance."

One perspective on medical problem solving generally can be characterized by Schmidt, Norman, and Boshuizen's (1990) proposed stage theory of clinical problem solving. They suggested that clinicians employ four stages of reasoning, which build upon, but do not replace, each other. Stage 1 involves the development of elaborate causal knowledge networks. Stage 2 involves compiling the elaborate networks into abridged ones. In stage 3, the clinicians develop illness "scripts," and in stage 4, they develop and use "instance scripts." Causal networks in the first two stages are based on underlying knowledge of pathophysiology, although stage 2 networks also are informed by observations of real patients. Stages 3 and 4 also are knowledge networks, but they are built increasingly on list-like structures containing information from real cases in the clinician's memory. The progression from one stage to another occurs as practitioners gain practical experiences that contribute to the cognitive representations of their knowledge. Causal networks of pathophysiology are not lost, nor do they grow as patient scripts slowly increase. The result is that most clinical problem solving ends up being script-based, rather than pathophysiology-based. Clinicians will tend to use the "highest" stage of reasoning available to them, depending on the availability and complexity of knowledge they have in any given domain. A clinician who has experience with a previous patient that fits a current patient will tend to use that experience as an instance script, whereas in an unknown situation the physician will revert to earlier stages until finding one that best accounts for the current problem. This theory explains the behavior of expert clinicians (and seems consistent with some other work done by researchers in the field of medical education (G. J. Groen & Patel, 1985; Guy J. Groen & Patel, 1991; Patel, Groen, & Norman, 1991) and with the generally accepted concept of the automaticity of expertise (Speelman, 1998; Winn & Snyder, 1996)). However, note that scripted behaviors are unlikely to entail problem solving as we have defined it, because those are the behaviors that experts use when addressing problems they are familiar with (or feel they are familiar with), and not when they address previously unencountered problems. Therefore, behaviors in "lower" stages are more likely to be characteristic of problem solving than are behaviors "higher" in the stages.

Research by Bordage and Lemieux (1991) sheds additional light on the model proposed above. They examined the semantic structures of experts and novices as such structures relate to diagnostic thinking. Their study, based on structural semantic theory, provided clinical problems for a variety of experts and novices and examined their solutions in terms of semantic relationships. These semantic relationships were used to reveal connections between the concepts that were relevant to solving the clinical problem. Each relationship identified was referred to as a "semantic axis." It was found that "the more diversified the diagnostician's network of semantic axes, the better was his or her diagnostic accuracy." In other words, the greater the number of valid abstract relationships the subjects found between bits of information in the problem, the more accurate were their diagnoses. The number of semantic relationships for each case was not necessarily related to the number of words used to describe the analysis. Therefore, while an expert might appear to arrive at a diagnosis quickly and without much thought (i.e., with very few words), a semantic analysis would reveal that the expert has understood and identified a great number of valid relationships between the data without spelling them out as such. Bordage (1994) later examined his and Lemieux's findings in terms of their implications for arriving at accurate diagnoses. In his words, "The most accurate diagnosticians, whether students or specialists, are those who have the most diversified sets of semantic axes (those who have elaborated or compiled structures) and who organize symptoms an signs into coherent systems of relationships of abstract qualities, and thus demonstrate a broader and deeper representation and understanding of the problem (p. 884)." Several related studies (Cholowski & Chan, 1992; Stevens, 1991; Stevens, Ikeda, Casillas, Palacio-Cayetano, & Clyman, 1999) appear to support Bordage's assertion. While we are reluctant to imply a false dichotomy between the perspectives proposed by these researchers, Bordage's statement, in comparison to Schmidt et al.'s (1990) assertion that "expert clinical reasoning is, to a large extent, based on the similarity between the presenting situation and some previous patient available from memory" (p. 617), highlights one of the fundamental tensions that has characterized this body of literature. We feel that the evidence now available, including our own research, tends to support Bordage's implication that to truly become expert, students must become expert at associating relevant information within a case, rather than or in addition to comparing cases. We will see that this approach, as embodied in the DP, is associated by our students with an increase in their ability to solve diagnostic problems.

Cognitive Tools

The term *cognitive tool* has been used in a number of related ways over the past decade and a half (see, for example, (Jonassen, 2003; Lajoie, Azevedo, & Fleiszer, 1998; Robinson, 1999; Salomon, 1988). Salomon (1988) presented the cognitive tool as a computer-based tool using or modeling an expert approach to a given process. He referred to the learner's interaction with cognitive tools as "AI in reverse." Whereas the goal of traditional artificial intelligence (AI) is to lead a computer to emulate the cognitive processes of people, Salomon suggested that learners' interactions with specifically designed software tools would lead them to acquire cognitive skills or strategies embodied in the software. Jonassen (2003) expanded the cognitive tool concept as follows:

> Cognitive tools are any technologies that engage and facilitate specific cognitive activities. They amplify the learners' thinking by enabling learners to represent what they know using different representational formalisms. As knowledge representation formalisms, cognitive tools are premised on the idea that humans learn more from constructing and justifying their own models of systems than from studying someone else's (p. 372).

In addition to the intuitive benefit of making one's own thinking apparent, Jonassen points out that that such a tool should also reduce the significant demand that complex problem solving places on working memory by harnessing the computer's ability to remember and organize. By this definition, the Diagnostic Pathfinder (DP) is a cognitive tool. The DP supports the learning of problem solving by presenting students with problems and requiring them to address those problems within the framework of the software. This paper does not accommodate a detailed description of how the Diagnostic Pathfinder functions. Such a description is provided elsewhere (Danielson et al., 2003). In brief, through a series of interactions, the DP presents a patient case that includes history, signalment, physical exam, and laboratory data, and then requires students to identify all abnormal laboratory data and communicate their diagnostic reasoning by organizing those data into a diagnostic path. The diagnostic path displays the students' diagnostic reasoning in an outline form of propositions that relate changes in laboratory data to the corresponding disease or physiologic processes occurring in the patient. After students commit to a diagnosis, a diagnostic path created by a faculty clinical pathologist is revealed for immediate comparison. This comparison allows both students and faculty to see the rationale used by the other when analyzing a patient's laboratory data.

Methods and results

Subjects

Between the Spring of 2002 and the Fall of 2004, the DP was used to teach eight semesters of Clinical

Pathology at five colleges of veterinary medicine. A total of 640 students participated in these classes, with roughly 70% of those students being female and 30% male. Other than one small pilot course of five students, the smallest class contained 42 students, with the largest containing 126, and the average class size being 80 students. All students participating in these courses were asked to complete a questionnaire regarding their experience with the DP software. Five hundred and forty three students completed the questionnaire. By class, this response rate varied from 46% (the lowest) to 100% (the highest), with the overall response rate being 84%.

Procedures

Students at the participating institutions used the DP to complete case-based homework assignments and prepare for exams. The number of DP cases assigned to each student using the DP varied from institution to institution, with the smallest number being 6, and the greatest number being 93. Curricular approaches at the institutions varied as well, though all but one of the institutions employed what might loosely be described as a traditional medical curriculum in which students received lectures in Clinical Pathology interspersed with laboratories involving the discussion of laboratory data for specific animal cases. The other institution, which only represents five of the students surveyed, uses a curriculum that mixes several approaches, including traditional strategies, collaborative learning, and problem based learning (PBL).

Instruments

The full questionnaire upon which the findings for this study are based is found elsewhere (Danielson et al., 2003). The questionnaire was designed to determine the students' perceptions of the DP's clarity (or usability), feasibility, and impact on learning. Responses to the items dealing with feasibility and clarity have changed systematically over time as the software has been debugged and various changes have been made in navigation, etc. We will not explore those responses in this paper. Rather we will focus specifically on the items designed to reveal the students' perceived impact of the DP on learning. Because the software's core learning interaction has remained largely unchanged, a comprehensive analysis of the students' reaction to the instructional attributes of the software can be performed meaningfully. The questionnaire items having to do specifically with learning outcomes are: 6, 7, 12, 16, and 17, and are found in Table 1. Other items, particularly those intended to measure enjoyment or ease of use, can arguably be hypothesized to indicate learning gains as well, at least indirectly. However, these items are also closely tied to software feasibility factors, such as computer bugs, network problems etc., so we will not discuss them here. In addition to the results of items 6, 7, 12, 16 and 17, we will examine responses to the survey's open-ended questions, many of which clarify the students' general indications of their response to the software's affect on learning. Those questions are as follows:

23. For questions above that you ranked particularly negatively, please indicate why here.
24. What are the things you like most about using the DP?
25. What are the things you like least about using the DP?
26. What would you change about the DP if you could?
27. Any additional comments you'd like to make about the DP:
28. If you used the DP for less than 20% of your cases, why did you choose not to use it?

Data analysis procedures

Descriptive statistics were calculated for responses to survey Likert items across all respondents by institution. Open-ended responses were analyzed initially by one of the primary researchers to reveal broad trends in the data. The responses were then codified according to those trends, and recorded in an access database by a research assistant. The coding was then reviewed and corrected as necessary by one of the primary researchers.

Presentation and Analysis: Likert Items

Table 1 reports student responses to the Likert items by institution and year. Students at all institutions generally indicated that the knowledge or behaviors identified in items 6, 7, 12, 16, and 17 were enhanced by DP use. Item 6 was intended to measure perceived completeness – i.e., how many data abnormalities students accounted for. Item 7 was intended to measure whether or not students felt that DP-use affected the precision of their diagnostic rationale, and, if so, whether the effect was positive or negative. As seen, students generally felt that DP-use made their rationale more precise. Items 12 and 17 were intended to function as general indicators of the students' appraisal of the DP's overall value to them as a learning tool. As seen by responses to those items, students generally found the learning value to be quite high. Finally, item 16 was intended to get a general sense of the DP's affect on the students' ability to organize the data relevant to solving a particular case. Again, as seen, the students generally indicated that the DP's affect on their ability to organize data was positive.

Table 1 *Mean responses by item number, institution and year:*

Institution/Yr	Mean	Median	Max	Min	Sample Size
6. Using the DP made me account for more lab data than I otherwise would have accounted for. less same more 1 2 3 4 5 6 7 8 9 10 NA					
Virginia Tech 2002	8.7	9	10	5	86
Iowa State 2002	8.5	9	10	2	96
Iowa State 2003	8.4	9	10	5	68
Wisconsin 2003	8.1	8	10	5	42
Wisconsin 2004	8.2	8	10	5	65
California Davis 2003	7.7	8	10	1	55
California Davis 2004	8.6	9	10	2	124
Guelph 2003	8.6	9	9	8	5
7. Using the DP made my diagnostic paths more precise than they would have been otherwise. less same more 1 2 3 4 5 6 7 8 9 10 NA					
Virginia Tech 2002	8.7	9	10	4	86
Iowa State 2002	8.1	9	10	1	96
Iowa State 2003	8.3	9	10	1	68
Wisconsin 2003	7.8	8	10	3	42
Wisconsin 2004	8.3	8	10	3	65
California Davis 2003	7.2	7	10	3	55
California Davis 2004	7.8	8	10	2	124
Guelph 2003	8.0	8	10	6	5
12. The DP makes doing my Clinical Pathology homework more worthwhile than similar paper-based assignments. definitely not absolutely 1 2 3 4 5 6 7 8 9 10 NA					
Virginia Tech 2002	8.6	9	10	1	86
Iowa State 2002	7.9	9	10	1	96
Iowa State 2003	8.6	9	10	3	68
Wisconsin 2003	6.8	7	10	1	42
Wisconsin 2004	7.4	8	10	2	65
California Davis 2003	6.1	6	10	1	55
California Davis 2004	7.3	8	10	1	124
Guelph 2003	8.8	9	10	7	5
16. Using the DP helps me to organize my thoughts about a case.					

	definitely not absolutely 1 2 3 4 5 6 7 8 9 10 NA				
Virginia Tech 2002	8.4	9	10	2	86
Iowa State 2002	8	9	10	1	96
Iowa State 2003	8.6	9	10	4	68
Wisconsin 2003	8.3	9	10	3	42
Wisconsin 2004	8.2	9	10	2	65
California Davis 2003	6.8	7	10	1	55
California Davis 2004	7.8	8	10	1	124
Guelph 2003	8.8	9	10	8	5

17. Using the DP makes understanding clinical pathology….

	harder easier 1 2 3 4 5 6 7 8 9 10 NA				
Virginia Tech 2002	8.3	9	10	3	86
Iowa State 2002	8.2	9	10	1	96
Iowa State 2003	8.9	9	10	6	68
Wisconsin 2003	8.5	9	10	4	42
Wisconsin 2004	8.4	8.5	10	5	65
California Davis 2003	6.9	7	10	3	55
California Davis 2004	7.9	8	10	1	124
Guelph 2003	8.3	8.5	10	6	5

Presentation and Analysis: Open-Ended Responses

We analyzed the responses to open-ended items with the hopes of understanding more specifically, from the learners' perspective, what characteristics of the software produced the results we had measured empirically in earlier studies. Here we will not report all open-ended responses, because of the large amount of information involved. Also, we will not report the results by survey item number, because there isn't always a predictable relationship between the nature of the question that was asked and the actual response. For example, the question, "What are the things you like most about using the DP?" elicited both of the following responses: "It was great in getting concepts stuck in my head. . . " and "I really didn't like it."

The open-ended responses were analyzed as follows. One of the primary researchers read through all the responses several times, identifying broad thematic categories. All the open-ended response data were then reviewed, coded, and categorized by a research assistant, in consultation with one of the primary researchers and with the aid of an Access database. This process resulted in several categories being combined, expanded, or eliminated. The resulting categories, in order of most responses to fewest, were as follows: 1. General Response, 2. Ease/Efficiency of Thinking 3. Ease of Use/Convenience, 4. Requirement that all data abnormalities be typed by hand and spelled correctly. 5. Requirement of completeness 6. Expert feedback, 7. Process of Manipulating Data in the diagnostic path, and 8. Diagnostic path format. Responses in each category were coded as either (a) positive comments, (b) negative comments and/or suggested improvements, or (c) comments that were mixed (both positive and negative) or in some way unclear. We then counted each response-type by category. Recall that all 543 students responded to the questionnaire. Because the Likert responses were largely positive, it would be expected that responses to open-ended questions would be as well. At the same time, because more open-ended questions were designed to elicit critical responses or suggestions for improvement (items 23, 25, 26 & 28) than positive responses (item 24), it seems reasonable to expect a disproportionately high number of critical comments.

1. General Response Category:

This category was used to broadly characterize the overall tenor of individual respondents' appraisal of the DP as a learning tool, considering all the responses to the open-ended questions. This category was considered important because many respondents gave mixed feedback (i.e., some positive comments and some criticisms or suggestions for improvement). All the comments for each respondent were considered cumulatively to determine if the overall impression of the software was positive or negative. In some cases, it was too difficult to identify one overall sentiment from a respondent's open-ended responses. Such respondents were classified as "unclear_neither_both". Of the 543 total respondents, 242 made comments that were judged to be predominantly positive, 152 made comments that were classified as unclear_neither_both, and 20 made comments that were predominantly negative. Many of the responses that contributed to this broad categorization will be discussed when we discuss the other nine specific response categories.

2. Ease/Efficiency of Thinking

A number of open-ended responses referred to the software's general effect on the way students perceive themselves as thinking about the problems. Of comments in this category, 178 were positive, 2 were mixed, and 5 were negative. The positive comments ranged from fairly vague statements such as, "a great way to get us to think clinically. . ." to more specific statements such as, "helps me be organized about my thoughts, shows me clearly flaws in my logic and ability to reason," and ". . . I could organize my thoughts in a logical manner." The two respondents who gave mixed responses in this category seemed to simultaneously feel that the DP was useful, but that it didn't change the way they would think about clinical pathology, or that it was inconsistent with their way of thinking. One respondent for example, made both of the following statements in response to different questions: "The DP did not make me understand clinical pathology any better. I still needed to use paper at times to organize my thoughts and group findings." and "[The DP is] very repetitive and good for learning how to rank and place abnormalities." Five students made comments suggesting that they felt the DP did little to help their thinking, or that the amount of effort required to use the DP did not justify the benefits that were derived from its use. For example, one respondent said, "I feel like you can breeze through the case without really learning a link between cause and effect." Another said, "I spent much more time using DP or putting this stuff on computer than it took to understand/make my problem list and reasonings on the lab data sheets. Not saying it probably didn't help cement things, but it was more time than necessary for my understanding (I felt)."

3. Ease of Use/Convenience

A number of respondents referred to the convenience or ease of using the DP. One hundred and fifteen respondents made positive comments in this category, 13 made mixed comments, and 21 made primarily negative comments. Some positive comments referred to specific aspects of the software that made study more convenient, such as the ability to work from any computer, the ability to work/save cases on-line, or the ability to save a partially completed case. Such comments included "online and total select" (the latter referring to the ability to select multiple data items in manipulating the data), and "being on the computer." Other positive comments were simple generic statements of convenience, such as "saves time", or "easy." Mixed comments often had to do with student responses to different aspects of the software, for example, one student objected to the fact that the software tied up the phone line, and experienced some trouble installing it, but still indicated that the software "Saves time, its easy." Other students indicated that their perception of the ease of use changed over time. For example, one said, "The first few cases seemed more overwhelming as the process was entirely new; once I was familiar with the program, there were no difficulties stemming from the program." The negative comments were usually tied to factors inherent in working with a computer. Some students did not own an adequate computer, and so completed their assignments on lab computers, which they did not find convenient. Other students expressed a preference for working without computers.

4. Requirement that all data abnormalities be typed by hand and/or spelled correctly

The DP requires that students manually enter names of all data abnormalities spelled correctly. Forty three students only made positive comments about this requirement, 10 students made mixed comments, and 71 students made primarily negative comments. Characteristic of positive comments, one student, when asked what he/she liked best about the DP wrote: "The structure of having to learn vocabulary by retyping. . ." Other students made only positive comments about this requirement, but only mentioned partial aspects of the process. For example, a number of students mentioned that they liked being told the correct spelling of data abnormalities after three tries. Most comments in this category emphasized that the primary benefit of this requirement was learning the vocabulary. One student providing a mixed review listed what he/she liked most about the program as "Learning new vocabulary," and what he/she liked least as "Having to repeat over and over the same vocabulary." Most of the negative comments in this category had to do with the requirement of typing each data abnormality name multiple times. For example, one student wrote, "Typing in lab data gets repetitive and annoying," and another wrote, "Sometimes just filling out the names of all the abnormalities was very time consuming." However, it seemed clear from many of these responses that it was the repetitiveness, and not the basic requirement of generating the names that most objected to. For example, the same student who made the previous comment also suggested, "Perhaps as one becomes more advanced, lab abnormalities can be pre-identified, making it easier/quicker to complete a case," suggesting that the requirement to type abnormalities is an acceptable entry-level requirement. Other students seemed to object to the requirement of correct spelling at all. For example, one student wrote, "I'd have a glossary for tests and abnormalities just for "pop up" spelling. I get mad having to retype for spelling," and another wrote that the software would be improved by "giving a break on incorrect spelling by 1 or 2 letters."

5. Requirement of completeness

Ninety seven students commented on the aspects of the software that require the learner to consider all laboratory data when constructing a diagnostic rationale, or to classify all data as being normal or not. Of those comments, 73 were positive, 9 were mixed, and 15 were primarily negative. Positive comments included, "couldn't ignore any abnormalities," "made me account for the data abnormalities," and "It made me analyze each and every piece of data, something I probably normally would not have done." One of the students providing a mixed review said, "The *DP* made me get more lab results than I probably would have gotten otherwise, but I sometimes found myself sifting thru them in a rote manner. Not really a *DP* problem-more my problem." Another wrote that what she/he liked best about the DP was that "It makes you account for every abnormality even insignificant ones.", while what she/he liked least about the DP was that, "It makes you account for every abnormality even on cases where the solution is obvious." One student providing primarily negative comments reported liking least, "Having to account for every extraneous, insignificant detail that['s] outside the norm." Another reported that she/he did not like, "Having to account for morphology results that were normal. No way to avoid putting them in the path."

6. Process of manipulating data in the diagnostic path/format of diagnostic path

Eighty nine students commented on the process of manipulating data in the diagnostic path. Forty six of these students made generally positive comments about this process, 4 made mixed comments, and 39 made primarily negative comments, or suggestions for improvement. Many of the positive comments had to do with the ease of manipulating the data. For example, when asked what they liked best about using the DP one student wrote, "I can move things easier than erasing them," another wrote, "neatness and ability to quickly and easily rearrange *diagnostic paths*," and a third wrote, "organizing clinical abnormalities and formulating a diagnostic path." One of the students providing a mixed review observed as a benefit that the DP, "helps me to analyze data more efficiently," while also noting as something she/he didn't like that it "takes a long time to construct a path." Many of the negative comments in this category had to do with the difficulty inherent in considering/presenting all the relevant data at once. For example, one student wrote that, "Not being able to see the entire path at once makes it difficult to organize my thought further along the path," and another observed that "It was difficult to tell what signs/data still needed to be placed into the diagnostic path." Other students had difficulty with specific aspects of the diagnostic path construction process, such as not being able to easily place mechanisms where they wanted them to appear in the diagnostic path. Finally, some students recommended that concept map-style diagrams be used in portraying the diagnostic path, as opposed to the outline format currently used.

7. Expert feedback

Eighty three students commented on the feedback they receive regarding their rationale in the form of the expert diagnostic path. Sixty eight of these comments were strictly positive; 15 were mixed. The following are illustrative of the positive comments: "I can compare my list to the expert list right away," "The professional pathways given very quickly," and "I liked having immediate feedback; I think that is very beneficial to learning something new." The fifteen students giving mixed comments all found the expert feedback useful, but wanted that feedback altered or expanded in some way. For example, one student wrote, "I wish the expert path had a few more notes explaining certain things instead of just listing them," and another wrote, "I would like it more interactive at the end when I compare my diagnosis with the clinical pathologist's." Several students also indicated wanting access to the expert rationale without having to complete the case first. None of the comments suggested that students did not want expert feedback.

Discussion

The Likert responses provide a valuable perspective for interpreting the open-ended responses. Clearly, based on the Likert response data, most of the negative comments or suggestions for improvement came from students who overall felt the DP was a useful learning tool. This is supported by findings from comments in the first category. Overall, positive comments outnumbered mixed comments by a ratio of 1.6 to 1, and positive comments outnumbered negative ones approximately 12 to 1.

The greatest number of positive comments fell into the second category, which we have called "ease/efficiency of thinking" because these comments tend to make claims regarding the DP's effect of helping to organize or clarify thoughts, or to make learning/thinking easier. These comments are perhaps the most difficult to interpret because this category contains the most vague positive statements. However, at the risk of stating the obvious, all these comments must refer to one or more things that the DP does, or allows. Also, since the same broad group of respondents produced comments in the other categories we have provided, it seems likely that many of the vague positive impressions, if the students could have been prompted for specifics, would have fallen into one of the broad categories 3 – 7. Similarly, because so many of these comments seem to center on the organization of thinking, it seems fair to say that these indications support the broad claims of cognitive tools enthusiasts that encouraging students to create representations of the knowledge relevant to solving a given problem will enhance their ability to solve that (and similar) problems. Furthermore, the fact that the enforced organization of pathophysiologic concepts is identified by the students as helpful to learning, and that it is associated with greater performance on related problem solving tasks, supports the idea that improved problem solving performance is related to practice producing robust representations of the problem's sub-elements.

As seen with the third category, one hundred and fifteen students specifically mentioned finding DP-use convenient/easy. The main idea seems to be that many students consider it easier to deal with/manipulate information electronically than it would be in paper form. This is true both in the context of the software itself, where concepts can be identified and manipulated without the necessity of erasing, re-writing, etc., as well as in the administrative aspects of the homework process (receiving assignments, partially completing and saving assignments, submitting assignments, etc.). Clearly, the data manipulation aspects of this preference could be theorized to have a beneficial impact on learning. It certainly seems reasonable to suggest that any relief an instructional approach may pose in administrative cognitive overhead will allow learners to focus more attention on integrating/understanding domain-specific concepts.

The fourth category, the one requiring students to enter all data abnormality names manually and spelled correctly, received by far the greatest proportion of negative comments. In fact, the number of students who felt this policy required a change outnumbered those who made positive comments about it by a ratio of 1.7 to 1. At the same time, among the forty three students who endorsed this feature, as well as among many arguing for change, there was strong agreement that writing data abnormality names manually resulted in learning the vocabulary. This was the intended outcome of this feature. It will be our goal to adjust the current requirements to the point that student learning and annoyance are optimized against each other. We plan to do this by allowing students to "pass out" of the spelling requirement by spelling data abnormality names correctly the first few times they encounter them, and then either having an auto-fill feature or a pull-down list become available for those students.

The fifth category referred to various gating behaviors of the DP that require students to consider all data that technically fall outside of the reference ranges for every test in any given case (even if such data turns out to be clinically insignificant). This feature was viewed as beneficial for learning by many students, with positive comments outnumbering negative comments or suggestions for improvement by a ratio of 4.9 to 1. In essence, many students felt that they considered more information than they would have considered had they not used the DP, and that doing so resulted in superior understanding of the underlying pathology and physiology. While a reasonable fear might be that this aspect of the software would serve as a crutch rather than a scaffold, with students returning to "ignoring" behavior with the DP withdrawn, that does not seem to have been the case, given the students' improvement on case-based exams that were entirely paper based.

The sixth category involves the process of manipulating data in the diagnostic path. Recall that this is a process of dragging and dropping concepts or clusters of concepts in an outline format, where it is understood that items above and to the left cause items below and to the right, or items below and to the right are supportive of items above and to the left. Students having only positive things to say about this process slightly outnumbered those having primarily criticisms or suggestions for improvement (1.2 to 1). Clearly, this process worked well for many students, but also was difficult for a significant number of others. While some other formalism for representing the learners' knowledge might be less problematic for more students, there are reasons to suspect that this process will be difficult regardless of the formalism used. First, the information-synthesis task is very difficult, regardless of the mechanisms used to assist the learner in accomplishing it. In this case, students had to represent causal and/or supportive knowledge involving dozens of discreet concepts in most cases, and more than a hundred discreet concepts for complicated cases. Given the known limitations of human cognition, this process is inevitably difficult regardless of the information-presentation formalism that is employed. Second, current commonly available electronic display technology doesn't permit for the spatially meaningful representation of large numbers of concepts simultaneously. This number can be very low for students with older computers that only display at a resolution of 640X480 or 800X600 pixels. Therefore, some students would not be able to represent as many concepts using the DP as they might be able to do using other traditional mechanisms of data representation, such as a blackboard, or a large piece of paper. In summary, our future research/development will explore providing optional ways of representing the diagnostic rationale, such as concept maps, with the hope of maximizing clarity and spatial efficiency, and accommodating individual learner preferences. At the same time, however, we suspect that the complexity of the task coupled with display limitations will continue to be problematic for learners, regardless of the formalism used to display the data.

The data presented in the seventh category show significant support for the DP's feedback feature whereby students receive feedback regarding their rationale in the form of an "expert rationale" created by an expert clinical pathologist completing the identical case using the DP. All respondents who mentioned this feature felt that this feature was valuable, with some requests that it be expanded in some way. This should come as no surprise to most readers – immediate and meaningful feedback is prescribed by many common instructional approaches, and these data support the general idea that feedback is a good idea. Again, one particular strength of this approach is that the expert and the student both used the same process to construct the diagnostic rationale, making it possible to attribute differences between student and expert rationale to different understandings of the underlying concepts, rather than to differences in the representation of the problem.

Implications for the Design of Software Tools

One of the complexities of attempting to meaningfully synthesize open-ended responses is that structure and order have not been pre-imposed on the resulting data, as commonly occurs with more traditional empirical approaches to research. The results can frequently represent a hodgepodge of ideas. In this case, however, we feel that several strong coherent concepts emerged that can be useful to the designer of software learning tools. These ideas should be evident in the discussion section. Here we briefly review them. 1. Our findings appear to support the general cognitive tool concept – that requiring learners to create detailed representations of the relationships between the concepts required to solve a given problem will promote understanding and problem solving ability in that domain. 2. Meaningful feedback is important. In our context, producing and presenting feedback in the same manner used by students to create their rationale appears to have been beneficial. 3. Students saw benefit in the DP's effect on the organization of their thoughts. This was accomplished in two ways: first, through gating the tasks relevant to solving the problem (data identification first, followed by data synthesis), and second, through use of the outline-based diagnostic path formalism. 4. The ease/convenience of electronically manipulating problem elements, and managing the larger learning process was appreciated by many students, and may have decreased the "data management" cognitive requirements for students.

Limitations and Future Directions

This study is part of a larger research project. These Likert-based and open-ended response data are intended to explore the specific aspects of the DP that produced the learning gains that were seen. While these findings seem compelling, the fact that they are based on the subjective analysis of open-ended response data suggests further studies to explore these principles empirically. For example, a version of the DP could be created that relaxes gating requirements. Second, the problem-solving tasks in which the students engaged constitute complete and authentic diagnostic problem solving tasks from the perspective of clinical pathology. Often, all the data a clinical pathologist has to consider are laboratory data, signalment, and a brief history. However, this process is not representative of the broader clinical problem solving process, which involves additional data and data collection, as well as the handling of therapy. In the future, we must further explore outcomes of DP-use using broader measures of problem solving, and including clinical problem solving performance.

References

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.

Bodner, G. M. (1991). A View From Chemistry. In M. U. Smith (Ed.), Toward a Unified Theory of Problem Solving: Views from the Content Domains (pp. 21-33). Hillsdale, New Jersey: Lawrence Erlbaum, Associates.

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.

Cholowski, K. M., & Chan, L. K. S. (1992). Diagnostic reasoning among second-year nursing students. Journal of Advanced Nursing, 17(10), 1171-1181.

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., Bender, H. S., Mills, E. M., Vermeer, P. J., & Lockee, B. B. (2003). A Tool for Helping Veterinary Students Learn Diagnostic Problem Solving. Educational Technology Research and Development, 51(3), 63-81.

Ericsson, K. A. (2003). The Acquisition of expert performance as problem solving: Construction and modification of mediating mechanisms through deliberate practice. In J. E. Davidson & R. J. Sternberg (Eds.), The Psychology of Problem Solving (pp. 87-126): Cambridge University Press.

Forde, R. (1998). Competing conceptions of diagnostic reasoning--is there a way out? Theoretical medicine and bioethics., 19(1), 59-72.

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

Groen, G. J., & Patel, V. L. (1985). Medical problem-solving: some questionable assumptions. Medical education., 19(2), 95-100.

Groen, G. J., & Patel, V. L. (1991). A View From Medicine. In M. U. Smith (Ed.), Toward a Unified Theory of Problem Solving: Views from the Content Domains (pp. 1-19). Hillsdale, New Jersey: Lawrence Erlbaum, Associates.

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. Educational Technology Research and Development, 48(4), 63-85.

Jonassen, D. H. (2003). Using Cognitive Tools to Represent Problems. Journal of Research on Technology in Education, 35(3), 362-379.

Lajoie, S. P., Azevedo, R., & Fleiszer, D. M. (1998). Cognitive Tools for Assessment and Learning in a High Information Flow Environment. Journal of Educational Computing Research, 18(3), 205-235.

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.

Quesada, J., Kintsch, W., & Gomez, E. (in press). Complex Problem Solving: A field in search of a definition?

Robinson, K. (1999). Cognitive tools in interactive design for digital design. In K. Martin, N. Stanley & N. Davison (Eds.), Teaching in the Disciplines/Learning in Context (pp. 354-359). Perth, UWA: The Eighth Annual Teaching Learning Forum, The University of Western Australia.

Salomon, G. (1988). AI in reverse: Computer tools that turn cognitive. J. Educational Computing Research, 4(2), 123-139.

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

Smith, M. U. (1991a). A view from Biology. In M. U. Smith (Ed.), Toward a Unified Theory of Problem Solving: Views from the Content Domains (pp. 1-19). Hillsdale, New Jersey: Lawrence Erlbaum, Associates.

Smith, M. U. (Ed.). (1991b). Toward a unified theory of problem solving: Views from the content domains. Hillsdale, New Jersey: Lawrence Erlbaum.

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

Speelman, C. (1998). Implicit Expertise: Do we expect too much from our experts? In K. Kirsner, C. Speelman,

M. Maybery, A. O'Brien-Malone, M. Anderson & C. MacLeod (Eds.), Implicit and explicit mental processes (pp. 135-147). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Stevens, R. H. (1991). Search path mapping: A versatile approach for visualizing problem-solving behavior. Acad Med, 66(9 Suppl), S73-75.

Stevens, R. H., Ikeda, J., Casillas, A., Palacio-Cayetano, J., & Clyman, S. (1999). Artificial neural-based performance assessments. Computers in Human Behavior, 15, 295-313.

Wenke, D., & Frensch, P. A. (2003). Is Success or Failure at Solving Complex Problems Related to Intellectual Ability? In J. E. Davidson & R. J. Sternberg (Eds.), The Psychology of Problem Solving (pp. 87-126): Cambridge University Press.

Winn, W., & Snyder, D. (1996). Cognitive perspectives in psychology. In D. H. Jonassen (Ed.), Handbook of research for educational communications and technology (pp. 112-142). New York: Simon & Schuster Macmillan.