COGNITION AND INSTRUCTION, 22(1), 61–79

What's So Good About Problem-Based Learning?

Noel Capon
Graduate School of Business
Columbia University

Deanna Kuhn
Teachers College
Columbia University

In a systematically designed and controlled experiment conducted in a naturalistic instructional setting, we examined adult students' learning of two concepts. Two intact classes taught by the same instructor were assigned to 1 of 2 conditions. In 1 class, instruction was problem based for 1 concept. For a second concept, lecture/discussion was the exclusive method. In the other class, matching of concept and method (problem based or lecture/discussion) was reversed. Two forms of assessment of learning occurred 6 and 12 weeks following instruction. At the initial assessment, the lecture/discussion group showed superior learning for 1 concept and the groups performed equivalently for the other concept. At the later assessment, however, the 2 groups showed equivalent ability to access each of the concepts, but each group showed superior explanation of the concept for which they had experienced problem-based learning. Results support the hypothesis of integration of new information with existing knowledge structures activated by the problem-based experience as the mechanism by which problem-based learning produces its benefits.

Do students learn more effectively if their learning is situated in the context of problems they are asked to solve? This question is of such broad relevance to theorists of many persuasions as well as to practitioners and to educators from the preschool to the graduate school level that practitioners might expect to have evidence by now weighing on the side of one answer or another. Some practitioners see a positive answer as self-evident to anyone with experience as a teacher or even a

Requests for reprints should be sent to Noel Capon, Columbia University, Graduate School of Business, 501 Uris Hall, New York, NY 10027. E-mail: nc7@columbia.edu

learner and hence scarcely requiring documentation. To those who have sought systematic evidence on the matter, however, data have been scarce, and no simple conclusions are forthcoming even though interest in the topic remains strong (Evensen & Hmelo, 2000). Moreover, the wide variety of practices that have been regarded as exemplars of problem-based learning contribute to the challenge of conducting rigorous research on the topic.

The handful of studies that have attempted to examine the question under controlled experimental conditions have been viewed skeptically on the grounds of questionable relevance to real-world learning. Needham and Begg (1991), for example, presented undergraduates a series of logical "brain teasers" (e.g., how to distinguish a liar and truth teller asking only one question). Under one condition, participants were asked to try to solve the problem before being told the solution. Under another, they were asked to memorize the problem and then were told the solution. When subsequently presented a set of analogous problems, those who had tried to solve the original problems were more successful (although the other group remembered the original problems better).

Studies designed to investigate the contrast of concern here under more naturalistic conditions have been difficult to interpret. The largest set of studies (e.g., see Albanese & Mitchell, 1993, for review) addresses problem-based learning in the education of medical doctors. A number of medical school programs have introduced experimental curricula in which students encounter patients and undertake to make diagnoses (situation-based or problem-based learning) in the early years of their training. Such curricula contrast to the traditional medical school curriculum in which clinical work only occurs later in training following initial years devoted largely to lecture courses in basic science subjects.

A recent review (Colliver, 2000) of studies comparing these curricula to traditional ones concludes that no compelling evidence exists for the superiority of the problem-based curricula. The severe methodological problems that Colliver identified as plaguing these studies, however, make it clear that no definitive conclusions can be drawn. Most serious among these is weak specification and control of content and delivery differences between the two curricula being compared and lack of random assignment to the curricula, leaving open the possibility of preexisting differences between groups. In a response to Colliver's review, Norman and Schmidt (2000) suggested the need to remain focused for now on basic research that allows for systematic examination of the multiple variables involved in naturalistic studies.

Our goal of the research reported here was to conduct a study of problem-based learning that achieves an effective balance between the dual objectives of naturalism (and hence relevance) and experimental control. We conducted the study in the context of a genuine educational context (a graduate school course), yet critical variables are either controlled (the instructor) or systematically varied in an unconfounded design.

A further goal of our work was to demonstrate specific rather than more global or extended outcomes of problem-based learning. The latter has been the norm in research in this area, with investigators typically seeking to show that students who experience an extended problem-based curriculum exhibit to a greater degree than a control group individual characteristics such as self-monitoring, self-regulation, planning, a positive orientation toward learning, and satisfaction with the learning process (Albanese & Mitchell, 1993; Blumberg, 2000; Hmelo & Lin, 2000). In our view, it is desirable to first establish local effects of problem-based learning, that is, specific learning outcomes that occur in a circumscribed time frame and are clearly attributable to the instructional method. Among other benefits, this approach is likely to prove most illuminating with respect to understanding of the mechanisms involved in problem-based learning.

Our final goal of this work was to shed light on the question of mechanism, assuming that a positive effect of problem-based learning is demonstrated. The idea that activity-based learning is superior to more passive modes has been a staple of educational theory for some time, with origins as varied as the constructivism of Piaget or Dewey and modern cognitive science (Kolodner, Gray, & Fasse, 2000). The active mode is invariably promoted as the desirable alternative but often in such a wholesale way as to discourage more precise definition of its critical features or analysis of its presumed benefits (Bereiter, 2002). Dual predictions are nonetheless possible: Problem-based methods promote active engagement, but lecture methods allow more material to be covered, in particular the multiple and varied exemplars that have been associated with superior acquisition and transfer (Gick & Holyoak, 1983; Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Singley & Anderson, 1989; Sternberg, 1985).

The most common general proposal as to the mechanism by which problem-based learning achieves a positive effect is that the experience activates a mental model that the student entertains with regard to the problem at hand. Once activated, this model facilitates performance (Schmidt, DeVolder, DeGrave, Moust, & Patel, 1989; Schwartz & Bransford, 1998). Although as Schwartz and Bransford (1998) noted, multiple mechanisms may operate, this general formulation in our view leaves open three distinct possibilities. A mental model activated by engagement with a problem may produce

1. Superior acquisition of new material (because of previously activated knowledge structures to which it can be connected).
2. Superior recall of new material (due to an increased number of retrieval paths).
3. Superior integration of new material with existing knowledge structures (leading to restructuring and enhanced conceptual coherence).

This study was designed so as to have the potential to shed light on the relative correctness of these three possibilities.

The variables we systematically varied in this study are the concepts being learned and the instructional method. Participants were two intact classes of 60+ students enrolled in an Executive MBA program at a major business school. Both classes were taught by the same senior professor who had taught the course for many years using a mixture of case study (problem-based) and lecture/discussion methods. Students were thus familiar with both instructional methods and accustomed to their combined use in the course.

Two concepts from the course syllabus—economic value to the customer (EVC) and lifetime customer value to the firm (LTV)—were identified as largely independent of one another and each amenable to teaching by either a problem-based or lecture/discussion method. Both concepts were taught to each class in a single 2-hr, 45-min class session with EVC taught first followed by a break and LTV second. In one class, the instructor taught EVC using the problem-based method and LTV using the lecture method. In the other class, EVC was taught by the lecture method and LTV by the problem-based method. Students were not aware of the manipulation. Both classes took place on the same day. An advantage of this design, of course, is that any observed superiority of either instructional method will be specific to the concept being taught by that method.

Preexisting time and format constraints precluded our making any major changes to the manner in which students' mastery of these concepts would normally be assessed. Our interest was more in long-term than immediate mastery of the concepts. Would students be able to access the concepts at a later time and understand and apply them appropriately in contexts that called for them? We therefore chose the final course examination, scheduled for12 weeks following the instructional session, as the context for one of two types of assessment. At this 5-hr examination, students had access to all their books and notes. The question on this exam that was relevant to this study consisted of an open-ended essay based on only a general prompt, one that gave students the opportunity to access and apply either or both of the concepts involved in the study. Although books and notes could have provided cues to either concept, no more direct cues to either concept were provided.

As a second, more directly cued form of assessment, the instructor added an unannounced quiz midway between instruction and the final exam (6 weeks postinstruction). Its purpose was to assess whether processing of the two concepts as opposed to later access (i.e., recall) differed. The concept was therefore provided and the student asked to define and explain it without access to notes. Again, because our interest was in longer term rather than immediate mastery, this intermediate interval seemed most informative. (Arguably, it would have been desirable to assess immediate mastery as well as to administer both types of assessments at each of the two assessment intervals, but again practical considerations precluded this degree of infringement on instructional time.)

METHOD

Participants

Participants were 131 students, 69 in Section A and 62 in Section B, in their second term in an Executive MBA Program at a major business school in the Northeast United States. They represented 93 different corporations and had a minimum of 5 years of executive experience. Their ages ranged from 29 to 45 years, with a mean of 33 years; 73% were men; 13% were non—Americans. The program comprises 20 courses and is completed in five terms. Students are randomly divided into two sections for their core coursework. Students in this program are financed by their employers and are highly motivated to perform well.

Concepts

Both concepts rest on the more generic concept of value. Customers and firms have value to one another. One concept focuses on the value of the firm's product to the customer. The other focuses on the value of the customer to the firm.

EVC. *EVC* is defined as the maximum price the customer would be prepared to pay for the product relative to the next best competitive alternative. The EVC is calculated by adding the product's net lifetime savings (summing costs and benefits) to the purchase price of the next best alternative.

LTV. *LTV* is defined as the discounted gross margin earned from the customer during the life of the customer–firm relationship minus annual maintenance costs. Its use highlights the relative values of acquiring customers versus retaining customers and the costs of customer defection.

Procedure

Instructional conditions. The sequence of instructional segments that defined each of the conditions appears in Table 1. The discussion segment (DISC) is unique to the lecture/discussion condition, and the three problem-based (PB) segments are unique to the problem-based condition. Other segments are comparable in substance (although not detail or time) across conditions. Approximate times devoted to each segment appear following each segment.

Students were assigned two case studies to read for the class session. One case provides a context for developing the EVC concept and the other a context for developing the LTV concept, but neither case study explicitly mentioned either concept.

TABLE 1
Instruction Conditions

Problem-Based	*Lecture/Discussion*
INT: Introduction of topic[a] (3 min)	INT: Introduction of topic[a] (3 min)
PB1: Problem Segment 1; students address problem in small groups and report back[b] (7 min)	CNCPT: Introduction and illustration of concept[a] (9 min)
CNCPT: Introduction and illustration of concept[a] (9 min)	EX: New example of concept[c] (6 min)
PB2: Problem Segment 2; groups assigned task of utilizing concept to solve problem[b] (23 min)	DEM: Demonstration of concept utilization[c] (23 min)
PB3: Problem Segment 3; groups report[b] (15 min)	DISC: Discussion of more examples and related concepts[b] (25 min)
DEM: Demonstration of concept utilization[c] (7 min)	
EX: New example of concept[c] (4 min)	

Note. Number of minutes devoted to segment indicated in parentheses at end of each segment. [a]Segment in both conditions and identical time. [b]Segment unique to condition. [c]Segment occurs in both conditions

In the first problem-based segment, the problem was posed in a generic manner. For the EVC concept, for example, the problem posed was how to approach setting a price for a new product. The groups in this segment were 2 to 3 persons in size (students were asked to discuss with the person or persons sitting next to them).

The two case studies the students had been assigned provided the basis for segments PB2 and PB3. As is typical in business case studies, the problem was to decide among several alternatives which was the best option for a company to pursue. Doing so required use of the relevant concept (EVC or LTV). The groups in segments PB2 and PB3 were 4 to 6 persons in size. In keeping with normal classroom practice, these groups were not facilitated or monitored. As a result, no specific data are available regarding what occurred during this group activity. The general observation, however, documented on videotape, was that all groups were engaged and talkative during the time allocated for the activity.

In both conditions, the instructor distributed copies for students to keep of the overheads presented in class. The lecture/discussion condition consisted largely of the instructor's presentation accompanied by a slide presentation. Students had the opportunity to ask questions or make comments, but student input comprised only a small portion of the class time and largely involved questions of clarification. The Appendix contains a key excerpt from the concept segment (CNCPT; which occurs in both conditions) for the EVC concept in which the instructor presents the definition and use of the concept. Note that the instructional conditions did not differ on a concrete versus abstract dimension. In both conditions, instruction was situated in the context of concrete examples.

Directly cued assessment. Six weeks later, students were given an unannounced quiz that directly cued each concept (EVC and LTV). They were not allowed access to any notes or other materials. Only those students in attendance at the beginning of the class on this day completed the quiz—52 in Section A and 51 in Section B.

The quiz comprised two questions:

> You are called in to meet with one of your senior executives who knows that you are enrolled in an executive MBA program. "So," she says, "I assume that you have learned something about pricing. We are about to launch a new line of light bulbs. How should we think about pricing them?"
>
> A second senior executive says to you "When I took a marketing course a few years ago, the professor always used to talk about the notion of customers as assets. I found that all very interesting. Has there been any conceptual development in this regard that can help us in the electrical fixtures market?"

The term *pricing* served as a cue to the EVC concept and the term *customers as assets* served as a cue to the LTV concept. Each question was printed on a separate sheet, and the two unfastened sheets were handed out together, with one of the sheets on top in half of the cases and the other sheet on top in the other half. Students were given 30 min to complete both questions.

Indirectly cued assessment. This assessment took place 12 weeks following the session at which the concepts were introduced and 6 weeks following the cued-assessment quiz as one question on the open-book final examination for the course. The examination was 5 hr in length. Students were advised of the percent credit for this question. The question, designed to elicit both the EVC and LTV concepts, although neither was specifically mentioned, was as follows:

> Customers and firms have value to one another. What concepts are available to identify and assess this value? If any such concepts are quantifiable, indicate how.

RESULTS

Directly Cued Assessment

We devised coding systems for students' answers for each of the concepts based on examination of a portion of the responses. Each system was revised and then applied to the remaining responses by two coders who were blind to the student's

condition. One fourth of these responses were coded independently by the two coders for purposes of reliability calculation. For the EVC responses, percentage of agreement was 87%, and for the LTV responses, it was 90%. Differences were resolved by discussion.

EVC levels. Responses to the EVC question were assigned to one of six levels, defined as follows:

0. No relevant response. (The student fails to invoke the key construct of value to a customer and typically incorrectly attaches the EVC label to some other concept from the course.)
1. Value of the product to the customer is mentioned as a consideration, but this value is not specifically identified as monetary.
2. Monetary value of the product to the customer is mentioned as a consideration, for example, "the maximum amount the customer would be willing to pay" for the new product.
3. The EVC concept is identified (although not necessarily by name) as the maximum the customer would be willing to pay and is defined as the price of an existing available product plus the net of savings and costs of the new product.
4. The EVC concept is applied to pricing with the EVC identified as the top of the range of potential prices and the variable production cost of the product as the bottom of the range.
5. Implications of the EVC concept are derived with respect to choosing among options for pricing, for example, to gain a market toehold and low expected sales, price near the EVC and to maximize market share and sales, price near the variable production cost.

LTV levels. Responses to the LTV question were assigned to one of four levels, defined as follows:

0. No relevant response. (The student fails to invoke the key construct of value of a customer to the firm, typically incorrectly attaching the LTV label to some other concept from the course.)
1. Retaining existing customers is identified as a consideration and potential asset to the firm.
2. Retaining existing customers is distinguished from acquiring new customers and is identified as a potentially more valuable goal in the long run.
3. The LTV concept is identified (although not necessarily by name) as the monetary gain the firm anticipates from a customer during the entire course of that customer's relationship with the firm.

TABLE 2
Percentages of Students Performing at Each Level in the Directly Cued Assessment

	Condition	
Concept	*Lecture/Discussion*	*Problem-Based*
EVC		
EVC concept identified and applications or implications noted, or both	15	15
EVC concept identified	9	7
Monetary value concept identified	22	20
Value concept identified	15	17
No relevant response	39	41
LTV		
LTV concept identified	29	15
Customer retention and acquisition distinguished	24	17
Customer retention concept identified	16	25
No relevant response	31	42

Note. EVC = economic value to customer; LTV = lifetime customer value to the firm. For the EVC concept, percentages for the two highest levels described in the text are combined.

Performance by condition. Percentages of students assigned to each level for the EVC concept by condition are shown in Table 2. As seen there, for each concept, only about one fourth of students were able to specifically define, apply, or draw implications regarding the concept after the 6-week interval from the time the concept was introduced. For the EVC concept, there was negligible difference across conditions. For the LTV concept, however, a difference appeared by condition, favoring students taught by the lecture method. These students were slightly more likely to give a Level 2 response (see previously) or higher, $\chi^2(1, N = 103) = 4.20$, $p < .05$. Note also that they were almost twice as likely to identify the LTV concept. (Categories were collapsed in the statistical analysis to focus on the relevant distinction.)

Indirectly Cued Assessment

Responses to this question averaged about three paragraphs or 300 words, although they ranged from a single sentence to three pages. The procedure for analyzing responses was similar to that used in analyzing responses to the cued assessment, with the difference that the entire response was examined for purposes of identifying any passages in which the student invoked either the EVC or LTV concepts (either in substance or by name). Any such passages were then assigned to one or the other concept, and each set was coded for level of explication of the concept. Students rarely interspersed passages relevant to one concept with passages

relevant to the other. Rather, if both concepts were invoked, the student addressed one first and then the other.

We devised coding systems relevant to each concept based on examination of a portion of the responses. Each system was applied to the remaining responses by two coders who were blind to the student's condition. One fourth of these responses were coded independently by the two coders for purposes of reliability calculation. For the EVC responses, percentage agreement was 91%, and for the LTV responses, it was 88%. Differences were resolved by discussion.

Students' responses were assigned to one of four levels for each of the two concepts:

0. No reference: The concept is not invoked, either in substance or by name.
1. Reference: The concept is mentioned, without definition or elaboration.
2. Definition: The concept is mentioned and the formal definition provided in class is given.
3. Explanation: The student explicates the concept, going beyond the formal definition to exhibit understanding of its meaning by identifying what at least some of the terms mean and how they relate to one another or how the concept might be applied in a specific case.

Illustrations of student responses in the latter two categories for the EVC concept are presented in Table 3 and for the LTV concept in Table 4.

Performance by condition. Performance by condition is shown in Table 5. As seen there, percentages of students who achieved the level of explanation (Level 3), although a minority of each group, differed significantly by condition for each concept. Students instructed by the problem-based method group (Section B for the EVC concept and Section A for the LTV concept) were over twice as likely to do so as students instructed by the lecture/discussion method, $\chi^2(1, N = 131) = 4.52, p < .05$ for the EVC concept, and $\chi^2(1, N = 131) = 3.92, p < .05$ for the LTV concept. (Categories were collapsed in the statistical analysis to focus on the relevant distinction.) Students instructed by the lecture/discussion method, in contrast, were more likely to simply give the textbook definition, which they could do by searching their notes, assuming they have been successful in retrieving the concept from memory (which more than two thirds of students were across concepts and conditions; see Table 5).

DISCUSSION

What do these findings add to our knowledge of when and how problem-based learning achieves superior outcomes to more passive forms of instruction? The best existing discussions of the topic are ones based on systematic empirical data

TABLE 3
Examples of Student Responses for the EVC Concept

Definition

"A product's EVC is the price of the next best alternative plus the net lifetime savings of the new product."

Explanation

"On the other side of the table is the firm's value to the customer. The principal method available to identify and assess the value of a firm to the customer is EVC. EVC can be expressed by a formula: EVC = next best price of competitive product + net lifetime savings from new product. The first component of this equation, price of next best competitive product, is fairly straightforward. Assuming our new product is 'best,' what is the price of the next best product on the market? The second component is where the action is. The net lifetime savings from the new product quantifies (in dollars) all the benefits and costs to the consumer from using our new product as opposed to using the next best product. These benefits and costs are wide ranging and can include actual monetary costs and benefits as well as non-monetary costs and benefits perceived by the customer to the extent these can be quantified in dollar terms."

"An EVC analysis would look at the relative benefits a product has over another in terms of how much it will save in relation to another product. If you have a light bulb that lasts 1,000 hours and costs $2, versus a light bulb that lasts 10,000 hours and only costs $4, you can determine that the EVC of the second bulb is $16. You will have to replace the first bulb ten times at $2 apiece to get the same lighting time as the second bulb."

Note. EVC = economic value to the customer. In the last example, the student is incorrect in calculation of the EVC. If the second bulb lasts 10 times longer than a bulb priced at $2, the EVC is $20.

TABLE 4
Examples of Student Responses for the LTV Concept

Definition

"The LTV of a customer is the discounted gross margin from that customer net annual maintenance costs."

Explanation

"The value of a customer to a firm is a function of not only the number of customers a firm acquires but more importantly how many profitable customers they acquire and retain. This concept has led to a method of measuring the value of customers to a firm. The lifetime value of a customer is defined as the discounted gross margin earned from customer, net of the annual maintenance cost. For an infinite time horizon, LVC = $(m - a)/(1 + d - r)$, where m = constant gross margin, a = constant customer maintenance cost, and r = retention rate .... Firms can increase the LTV of their customers by increasing either the retention rate or gross margin, or by decreasing either the maintenance cost or the cost of capital."

"[Definition of concept, formula, and definition of terms provided as in preceding example.] In other words, lifetime value measures the discounted cash flows expected from the customer in the future as well as today, less the cost of maintaining the client. Since maintenance costs in period 1 are typically higher than in later periods because of acquisition costs, LTV therefore inherently measures the benefit of retaining clients over time. It proves the business adage that keeping a customer is usually more profitable than getting a new one."

Note. LTV = lifetime customer value to the firm. In the first example of the explanation level, the student omitted to note that d = cost of capital.

TABLE 5
Percentages of Students Performing at Each Level in the Indirectly Cued Assessment

Condition	*Explanation*	*Definition*	*Reference*	*No Reference*
EVC concept				
Lecture/discussion	10	49	14	26
Problem based	24	29	11	35
LTV concept				
Lecture/discussion	11	63	08	18
Problem based	25	43	04	28

on students' learning of meaningful academic subject matter (e.g., Patel, Kaufman, & Arocha, 2000; Schmidt et al., 1989; Schwartz & Bransford, 1998). These discussions go beyond the either/or debate as to which method is superior, especially as few proponents of the passive method are to be found. As Schwartz and Bransford (1998) highlighted in an incisive series of experiments examining the two methods, there is indeed a "time for telling." Didactic methods provide higher level explanation that would be time consuming, difficult, and even impossible for students to construct on their own. Schwartz and Bransford proposed, however, that an optimum time for telling is once students have discerned the features and structures that differentiate relevant aspects of the phenomena to be understood. Although Schwartz and Bransford did not engage students in the collaborative activity typical of problem-based learning, their findings support their thesis that superior processing of didactic material occurs when a student has first engaged in analyses of pertinent dimensions of the phenomena that are to be explained.

As in Schwartz and Bransford's (1998) work, the concepts being acquired in this study were not ones that students could have generated on their own in the absence of instruction. Social transmission of these concepts during the group process is therefore not a plausible explanation of superior understanding in the problem-based condition. Despite several years of business experience, students came to the class equally ignorant of the formal knowledge represented in these specialized concepts and a time for telling was clearly required if students were to acquire them. In contrast, in a study by Schmidt et al. (1989) that reported results like ours, students who first collaborated in small groups in addressing a problem related to the scientific concept of osmosis (why a red blood cell swells in pure water and shrinks in salt water) were compared with another group of students who discussed an unrelated problem. Students then read a text passage on osmosis and immediately following were asked to recall in writing all they remembered about the text. Although there was no difference in the number of descriptive statements, the groups who had previously discussed the topic produced significantly more explanatory statements. The fact that students already possessed varying degrees of

knowledge regarding the concept, however, makes it impossible to rule out social transmission of information during discussions as contributing to what students recalled. In this study, in contrast, social transmission is not an adequate explanatory mechanism.

In accounting for the superior performance they observed in problem-based conditions, Schwartz and Bransford (1998) distinguished a "knowledge assembly" view and a "discovery as discernment" view. In the former

> Analyzing the cases encourage(s) students to assemble relations that connect the case information to other pockets of prior knowledge. Conceivably, this elaboration increases the number of possible retrieval paths (connections) to the target concepts. The multiple retrieval paths increase(s) the chances of recovering the relevant concepts. (Schwartz & Bransford, 1998, p. 492)

In the discovery as discernment view, prior deliberation produces superior analysis and noting of patterns and hence better encoding and better later access. In Schwartz and Bransford's (1998) words, "individuals learn well when they have generatively discerned features and structures that differentiate relevant aspects of the world" (p. 493).

Schwartz and Bransford (1998) noted that these proposed mechanisms do not contradict one another and in fact most likely operate together. Still, the alternative explanations of mechanism we identified at the outset of this article remain. To what extent should superior explanation or understanding following problem-based learning be attributed to any or all of these three factors: (a) superior acquisition of new material, (b) superior recall of new material, or (c) superior integration of new material with existing knowledge structures?

Our findings shed some light on this question. The first assessment failed to support a hypothesis that the problem-based group would show superior acquisition or recall of the relevant lecture material. The lecture/discussion group showed as good or better representation of the concept at this point (as reflected in identification of the LTV concept; see Table 2). A possible interpretation of this difference is that the ECV concept has stronger intuitive roots, in particular, the everyday concept of what a customer is willing to pay; hence, when the formal ECV concept was introduced, students could more readily integrate it with existing knowledge structures. The LTV concept, having no such intuitive counterpart, did not offer this opportunity. Students in the LTV problem-based group thus suffered from the lack of repetition and varied examples that the LTV lecture/discussion group needed to make sense of the concept. For the more intuitively based ECV concept, in contrast, this repetition and illustration were not necessary and students performed as well without it as with it.

At the second assessment, the results show no difference across conditions in ability to produce the concepts. Two thirds or more of students in every condi-

tion were able, given the prompt "value," to at a minimum search their notes or overhead copies and produce each of the concepts (reference, definition, or explanation categories in Table 5), and half or more in each condition were at a minimum able to produce formal definitions. Where the difference between conditions arises is in the likelihood of students going beyond the definition of a concept to explicate its meaning or use (explanation category in Table 5). The probability of students doing so remains lower than might be hoped for both groups and both concepts, but a clear difference emerges between instructional conditions for both concepts.

In sum, then, of the three processes—acquisition, recall, and integration—postulated earlier, our data provide the strongest support for integration as the locus of differential effects of problem-based and traditional instruction. The best way to describe this effect, we believe, is to say that students who experienced problem-based instruction more often were able to integrate newly acquired concepts with existing knowledge structures that had been activated. In more everyday language, they demonstrated understanding. Ideally, we would like to have richer measures of integration than were feasible in this study. Would students draw on the concept when confronted with a relevant real-life situation? Nevertheless, the differences between responses to the second assessment categorized as definition and those categorized as explanation (illustrated in Tables 3 and 4) are clear enough, we believe, to warrant the inference that there was a significant, qualitative difference between the two types.

The benefit of problem-based learning we might tentatively conclude then lies not in superior acquisition or recall of new concepts but in the potential for greater understanding reflected in an integration of the new concept with existing knowledge, and with it, the possibility of restructuring and enhanced conceptual coherence. Put in simpler terms, the answer to what's good about problem-based learning is that it promotes sense making.

The research task now, as we see it, is to learn more about the cognitive mechanisms associated with the two kinds of learning contexts examined here and especially how to draw on them both in ways likely to optimize learning. Ideally, sequences of methods might be identified that lead to genuine and flexible understanding in a majority rather than only a minority of students and in less motivated students as well as the highly motivated ones examined here.

Also critical to a productive research agenda on this topic is more rigorous examination of what is and is not problem-based learning. As noted at the outset of this article, the term has been used to describe a wide variety of practices. Is social collaboration an essential component and if so, why? Students can certainly confront any number of problems in a solitary mode. Or, possibly, is the social stimulation necessary to ensure the desired level of cognitive engagement? Is the social exposure of having to present solutions to peers (or even expect to have to do so) a critical component? Is the analysis of alternative problem solutions (common to

the case method) sufficient to yield the cognitive benefits of problem-based learning or must students construct their own solutions? All of these questions are amenable to carefully controlled experimental investigation. In this work, we have examined only one form of problem-based learning in one type of student population. To conduct controlled investigation of different forms of problem-based learning in different populations is a demanding agenda, but the enormous effort educators are currently investing in problem-based learning initiatives suggests that it is an agenda worth the effort required.

REFERENCES

Albanese, M., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine, 68,* 52–81.

Bereiter, C. (2002). *Education and mind in the knowledge age.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Blumberg, P. (2000). Evaluating the evidence that problem-based learners are self-directed learners: A review of the literature. In D. Evensen & C. Hmelo (Eds.), *Problem-based learning : A research perspective on learning interactions* (pp. 199–226). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Colliver, J. (2000). Effectiveness of problem-based learning curricula: Research and theory. *Academic Medicine, 75,* 259–266.

Evensen, D., & Hmelo, C. (Eds.). (2000). *Problem-based learning: A research perspective on learning interactions.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gick, M., & Holyoak, K. (1983) Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Hmelo, C., & Lin, X. (2000). Becoming self-directed learners: Strategy development in problem-based learning. In D. Evensen & C. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions* (pp. 227–250). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kolodner, J., Gray, J., & Fasse, B. (2000). Promoting transfer through case-based reasoning: Rituals and practices in learning by design classrooms. *Cognitive Science Quarterly, 1,* 1–47.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Society for Research in Child Development Monographs, 60*(4, Serial No. 245).

Needham, D., & Begg, I. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory and Cognition, 19,* 543–557.

Norman, G., & Schmidt, H. (2000). Effectiveness of problem-based learning curricula: Theory, practice and paper darts. *Medical Education, 34,* 721–728.

Patel, V., Kaufman, D., & Arocha, J. (2000). Conceptual change in the biomedical and health sciences domain. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 329–391). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schmidt, H., DeVolder, M., DeGrave, W., Moust, J., & Patel, V. (1989). Explanatory models in the processing of science text: The role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology, 81,* 610–619.

Schwartz, D., & Bransford, J. (1998). A time for telling. *Cognition and Instruction, 16,* 475–522.

Singley, K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.

Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence.* Cambridge, England: Cambridge University Press.

APPENDIX
Excerpts From Segments Concept (CNCPT) and Demonstration (DEM; Problem-based Condition) for Economic Value to the Customer (EVC) Concept

Segment CNCPT

Instructor: One way to make progress in this economy is determining the price of a new product. What's the value we're delivering to a customer and what ultimately does that mean? That's the sort of world that we're in. Okay, the concept is economic value to a customer—something that's got economic value added, not to be confused with the financial economic value added. [See Figure A1 for overhead presented.]

So, the EVC is the maximum price the customer would be prepared to pay for the new product, considering the next best, competitive alternative. In other words, right now your customer is buying and using one particular product, product A. You come along with product B, which has some level of economic benefit. Presumably there's a point, all things being equal, at which the customer is indifferent between continuing to use product A and switching to product B.

Okay. Let's take a look at the tape. [Instructor presents a 5-min video summarizing the case study situation students had been assigned. The video depicts the customary method farmers use to reduce cannibalism among chickens, debeaking, compared to a new method in which the chicken are fitted with contact lenses, which achieve the same goal.]

- **Concept:** **The maximum price the customer would be prepared to pay for the new product considering the next best competitive alternative**

- **Definition:** **EVC = Purchase Price of the next best competitive alternative + the net lifetime savings from the new product**

- **Use:** **Helps to establishes the maximum value a customer would pay for the new product, *ceteris paribus***

FIGURE A1 Overhead 1: Economic value to (a) customer.

So, what does the farmer have to decide? He has to decide what he would save. He has to determine the EVC. [Segments PB2 and PB3 follow in the problem-based condition: Students are assigned to address problem in small groups and report back.]

Segment DEM

Instructor: So, to sum up, what types of savings might the farmer expect if he or she switches from debeaking, which is the reference point, that's what they're currently using, into the contact lens system?

Student: Less cost for food.

Instructor: What's the basis for that? How does that work? ... [pause] Okay, they eat more evenly; the food doesn't get kicked out onto the floor. What else?

Student: Fewer birds die.

Instructor: Okay. Do you remember the number? Does anyone remember the number of the mortality rate for non-debeaked? [Student responds.] No, that's the beaked. When somebody started doing the debeaking, they went from 25% down to 9%, which is where we are now, and we are headed to 4.5%. A 50% reduction in mortality, and what else?

Student: Labor costs?

Instructor: What about that? It seems almost a wash, isn't it, in terms of the ... presumably it's the same type of people who do debeaking will do lens insertion. It's a very similar sort of job. And what's the sort of throughput? In terms of the debeaking person vs. the lens person? It's about the same, right? I'm just making the assumption of skills that are probably pretty similar. So let's consider that to be a wash. Okay, what else?

Student: How many eggs the chicken lay.

Instructor: Yeah, okay. That's right. Okay. There's egg production. There's egg production. What else?

Student: What it cost them to make the lens.

Instructor: No, from the point of view of the farmer, he could care less about that. The cost of the lens to the manufacturer, ODI, represents the lowest point at which they could price. The cost of the lens to ODI is an irrelevancy for this part of the exercise. What we're trying to do, remember, what we're trying to do here. What's the cost to debeak, which is fact just the labor cost, right?

Okay, then, what's the value that the farmer would get if the farmer switched to the lens? [See Figure A2 for overhead pre-

sented.] Well, he'll gain some value by having less birds die [instructor indicates on overhead], some feed savings [indicates on overhead], some value for extra egg [indicates on overhead]. So the value he's going to get is going to be up here somewhere [See Figure A3 for overhead presented], right? And this is what we call the EVC, the economic value to the customer [instructor indicates upper line on overhead]. Right? That gives us a criterion, if you like, with which to figure what to price. We know that we cannot price above that number. We know that if we are right here, the farmer is going to be indifferent [instructor indicates upper line on overhead]. All things being equal, indifferent between debeaking and the lens. And the further we come down here [instructor indicates movement from upper to lower line], the more likely, the greater the value incentive, we are giving to the farmer to switch.

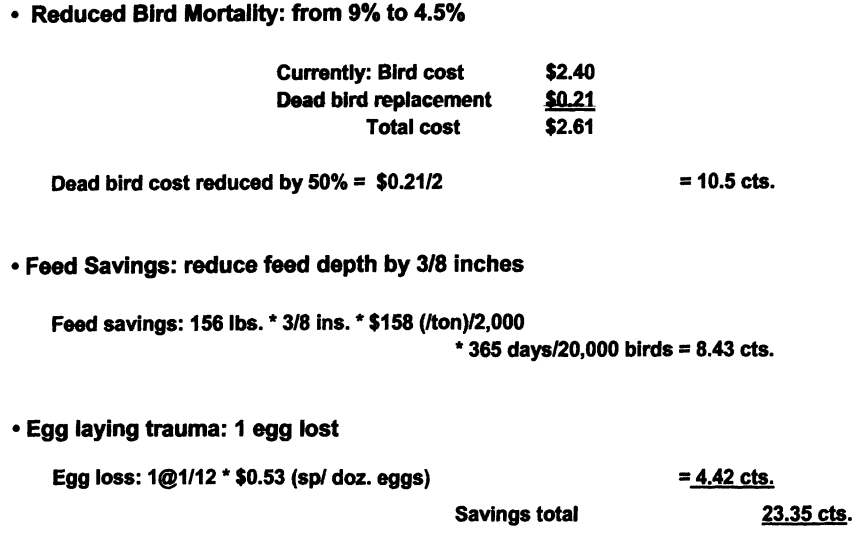

FIGURE A2 Overhead 2: Economic value added for manufacturer.

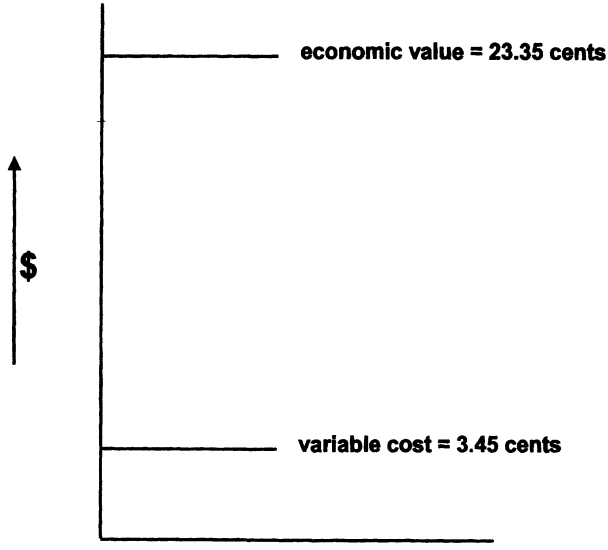

FIGURE A3 Overhead 3: How should the price be set?