Creating Problems for Teachers: Research on Constructing Problem-Based Materials to Enhance Science Content Knowledge

Joy Anne Oslund*, Mark R. Low*, Jamie N. Mikeska, Ayelet Weizman, Mary A. Lundeberg,

Matthew J. Koehler, and Janet Eberhardt

Michigan State University

This material is based upon work supported in part by the National Science Foundation, under special project number ESI - 0353406 as part of the Teacher Professional Continuum program.

Any opinion, finding, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of any of the supporting institutions.

* The contributions of the first two authors were equal.

Abstract

This study is an exploratory analysis of the characteristics of effective problem-based learning (PBL) materials in the context of the professional development (PD) of science teachers. The authors synthesize a literature on the writing of effective PBL materials, and they create a framework for evaluating such materials. They compare their analyses of four specific problems with data that they collected regarding the use of those four problems in a weeklong PD event. The authors argue that the data that they collected from conference participants and conference sessions corroborate their ratings of the PBL materials. The authors conclude that their framework is useful for evaluating PBL materials in the context of science teacher professional development.

Creating Problems for Teachers: Research on Constructing Problem-Based Materials to Enhance Science Content Knowledge

Objectives

This study is an exploratory analysis of the characteristics of effective problem-based learning (PBL) materials in the context of the professional development (PD) of science teachers. Although a body of literature exists regarding the characteristics of effective PBL materials, we did not find any research that reported empirical work, and we found little research that applies to the context of the professional development of teachers. We synthesized the literature on the writing of PBL problems (in many contexts) and cases (in education contexts), and we analyzed those characteristics of good problems in light of the data that we collected during a weeklong problem-based learning teacher professional development conference. In this paper, we first outline a theoretical framework for the use of problem-based learning materials, and we provide a synthesis of the literature on the writing of effective problems. Next, we discuss our methods and our analyses of the data that we collected during the weeklong professional development conference. Finally, we offer recommendations regarding the creation of problems for use with science teachers in professional development contexts.

Theoretical Framework

Contextualized-Learning Approaches

Many instructional approaches attempt to situate learning within a task that is contextualized or authentic for the learners. These approaches include problem-based learning, discovery learning, case-based education, and project-based instruction, among others. Two of these approaches, case-based and problem-based learning, incorporate realistic or factual narratives to provide opportunities for students to integrate multiple sources of information in authentic contexts, and instructors use these materials in an attempt to engage students in problems related to their discipline (Herreid, 1994; Lundeberg, Levin, & Harrington, 1999). We used the literature on both cased-based and problem-based materials because the literature pertaining solely to the development of PBL materials (i.e., the writing of good problems) is not substantial, and because many similarities exist between the two instructional approaches. In the following paragraphs of this section, we provide an overview of these two approaches, and we highlight some of their distinctive features.

Case-based instructional practices involve the use of narratives or videos of realistic or factual situations, but may or may not include a problem to be solved, and may or may not require participants to make a decision regarding the information provided. Further, they may or may not include all of the information needed to understand the case (we will contrast this with problem-based practices in the next paragraph). Cases may be used as exemplars of a certain concept or situation, or they may provide practice for knowledge and skills that have already been taught (for examples of cases, see Barnett, Goldstein, & Jackson, 1994; Greenwood & Fillmer, 1999). Although several instructors who use case-based learning methods have provided suggestions for how to write cases (e.g., Carter, 1999; Herreid, 1997; Shulman, 1996), few have empirically tested their assumptions regarding the characteristics of a good case.

In the PBL context, an instructor presents an ill-defined problem in a relevant context to a group of learners. The learners engage in collaborative inquiry, self-directed research, and problem-solving in an effort to make a decision about the problem. In problem-based learning, the instructor often presents the problem to the learners before the learners have had any deep experience with the subject matter contained in the problem. For example, in a problem about blood doping[1], an instructor would not first provide mini-lectures or class sessions on that topic. Rather, participants would need to rely on their prior knowledge of the circulatory system, as well as engage in collaborative inquiry and self-directed learning about blood doping in an effort to resolve the problem.

In short, the goals for learners participating in a problem-based learning process include: acquisition of content knowledge, development of their problem-solving skills, opportunities to collaborate with peers, enhancement of their self-directed learning skills (Chernobilsky, DaCosta, & Hmelo-Silver, 2004; Hmelo-Silver, 2004; Kelson, 2004), and fostering an intrinsic motivation to learn (Hmelo-Silver, 2004). A number of features of the PBL process distinguish it from other task-centered, contextualized instructional approaches (e.g., discovery learning, case-based education, project-based learning, etc.). For instance, these other instructional approaches may or may not begin with an ill-defined problem. In addition, in many contextualized approaches, the instructors teach the subject matter of the problem to the participants before it is presented in the problem (Albanese & Mitchell, 1993; Barrows & Tamblyn, 1980). This minimizes the need for participants to engage in self-directed learning (wherein participants find their own resources to gain information needed to make a decision about the problem); in comparison, self-directed learning represents one of the salient features of the PBL process (Albanese & Mitchell, 1993).

Problem-based learning has its roots in medical education and much of the research on its effectiveness comes from studies done in medical schools (Albanese & Mitchell, 1993; Barrows & Tamblyn, 1980; Dochy, Segers, Van den Bossche, & Gijbels, 2003). Albanese and Mitchell

---

[1] Blood doping is a performance-enhancing procedure done by athletes. It involves using a blood transfusion or medication to increase the amount of red blood cells in the athlete, which allows the cells to receive more oxygen and may improve endurance.

(1993), pioneers in the study of PBL in the medical school context, defined PBL as "an instructional method characterized by the use of patient problems as a context for students to learn problem-solving skills and acquire knowledge about the basic and clinical sciences” (p. 53). In this context, instructors provide students with written case histories of patients as a way of introducing the problem (Barrows & Tamblyn, 1980). After being given the patient case history, medical students (in groups of approximately 6) list cues or facts of the case, hypotheses, and learning issues (questions). After learning issues have been identified and prioritized, students are given time for research (self-directed learning). After a time, the group reconvenes to share information, agree on a diagnosis, and collaborate on an appropriate course of treatment for the patient.

While much of the literature on PBL resides in the context of medical schools, professionals in other fields have recently made efforts to adapt the process to their own contexts. For instance, problem-based learning has been used in nursing (Newman, 2001, 2004), undergraduate science (Allen, Duch, Groh, Watson, & White, 2003), teacher preparation (Butler, 2003; Derry, Seymour, Fassnancht, & Feltovich, 2001), business (Capon & Kuhn, 2004) educational psychology (Chernobilsky, DaCosta, & Hmelo-Silver, 2004), and middle-school science (Gordon, Rogers, Comfort, Gavula, & McGee, 2001). Much of the research on PBL outside of medical schools measures factors such as perceptions, beliefs, and satisfaction with the process. Research that examines learning outcomes is quite limited, but what does exist is promising. For example, Capon and Kuhn (2004) compared learning outcomes from a PBL approach versus learning outcomes from a lecture and discussion approach in an undergraduate economics course. They found that although both groups of students could provide rote definitions of the concepts, the students that experienced the PBL approach were able to go beyond definitions and more fully explain the concepts.

Despite the growing base of literature that describes the theoretical and practical use of problem-based learning, the amount of literature on the writing of problems for use in the PBL process is limited. In addition, because PBL has not been widely used with practicing teachers in professional development contexts, literature on the writing of problems for use with science teachers in such contexts is nonexistent. In the next section, we synthesize the literature that does exist on the writing of problems, drawing from both problem-based and case-based sources. In the "Discussion" section of this paper, we describe how we attempted to make sense of that literature in terms of the writing of problems for use in the professional development of science teachers.

Writing Problems

Several different sources list criteria for good case materials or good problem-based learning materials and, while the lists are not identical, many had several features in common. In synthesizing the characteristics of good problems, we developed a set of criteria for evaluating problems: authentic and relevant, appropriate in regards to curriculum payoff, controversial, appropriately complex, clear about participants' roles, and engaging. We present this framework in Table 1, and we elaborate upon our definitions for our criteria in this section. <Insert Table 1 about here.>

Authenticity and relevance. Authentic and relevant problems either entail professional situations that participants are likely to encounter or they describe situations of which participants are aware, and they pertain to current issues of the profession (Herreid, 1997). Authentic and relevant problems cause participants to empathize with the characters of the problem, and the decision-forcing moments of the problem are within the actual realm of powers possessed by the characters in the problem (Herreid, 1997). Similarly, authentic and relevant problems allow participants to play a role in the solution of the problem (Kelson, 2004).

Pedagogical effectiveness. Many instructional approaches exist regarding the teaching of science content, and each of these approaches vary with respect to their requisite inputs (for example, time and materials) and their expected outcomes. To assess the pedagogical effectiveness of a PBL problem is to make a judgment regarding the appropriate commitment of resources in its use. In other words, two considerations are made: first, whether a particular set of desired learning outcomes warrants the use of the PBL process, and second, whether the selected problem could potentially produce those learning outcomes. As we wrote above, the expected outcomes of the PBL process include: acquisition of content knowledge, enhancement of problem-solving skills, collaboration with peers, and opportunity for self-directed learning (Chernobilsky, DaCosta, & Hmelo-Silver, 2004; Hmelo-Silver, 2004; Kelson, 2004). Pedagogically effective problems warrant the cost in terms of teaching and learning time and effort, and they foster the development of the intended outcomes of the PBL process.

Controversial. Controversial problems provoke conflict and engender multiple, viable hypotheses about which reasonable people could disagree (Duch, 2001; Herreid, 1997; Kelson, 2004). Controversial problems may include various competing, and equally valid, solutions (Harrington, 1995). They are ill-structured, ambiguous, and can be interpreted in many useful ways (Hansen, 1997).

Appropriately complex. Appropriately-complex problems do not provide participants with an abundance of information, nor do they allow participants to rely on prior knowledge, but rather they require participants to locate information for themselves, using available sources

(Delisle, 1997). Appropriately-complex problems provide only the amount of information that participants would really possess if they were actually encountering the problem situation (Kelson, 2004). Such problems require participants to engage in an appropriate inquiry process in order to reach the most reasonable solution and justify their decision, and the self-directed research phase should require rigorous, academic work (Levin, 2001). Such problems necessitate the collaboration of all group members in efforts to solve them.

Clear about participants' roles. Problems that are clear about the roles of the participants have clear boundaries and naturally lead participants to a particular production or performance (Kelson, 2004). In other words, when participants read the problem, they are aware of what they are expected to produce by the end of the process. In addition, such problems specify the role of the learners as individuals who are stakeholders in the solution of that particular problem (Torp & Sage, 2002). Not only can participants empathize with the central character of the problem, they can also imagine themselves as the central character in the problem. A problem that is clear about the participants' roles requires participants to make a decision regarding the resolution of the problem, and the problem requires participants to justify their decisions.

Engaging. Engaging problems retain the attention of the participants throughout the PBL process, thereby promoting the intended outcomes of the process. Engaging problems "stimulate curiosity, arouse interest, and motivate participants to probe for deeper understanding" (Herreid, 1997; Kelson, 2004). When presented with an engaging problem, participants become active learners who take a personal interest in the dilemma; such problems drive participants to consider alternative solutions and pursue the information necessary to make the best decision possible.

Method

Participants and Professional Development Model

We studied a professional development (PD) event consisting of teams of teachers from K-12 science classrooms in an intermediate school district (ISD) in the Midwest. The PD event had two phases in 2005-2006, but for the purpose of this paper, we have chosen to study the PBL materials developed for the first phase only, in order to focus on materials that dealt strictly with science content. Forty-five teachers, 11 male (mean age, 37) and 34 female (mean age, 42), participated in the first week of the PD event, referred to as a *Professional Working Conference (PWC)*, which the problems we study were written for. Prior to the PD, these teachers each identified a unit or lesson that they were interested in working on during the course of the week. The primary activities for the first week included individual (and team) lesson/unit development time and PBL sessions designed to help develop content knowledge (for science topics in and out of their self-identified focus area). Of the 45 participants, 20 taught in self-contained elementary school classrooms, 20 taught subject-specific science classes in middle or high schools, and five were elementary specialists (for reading or science). None of the teachers had experienced problem-based learning before attending the *Professional Working Conference*.

During the *PWC*, facilitators divided teachers into three *mega-strand* groups (Earth/Space Science, Physical Science, and Life Science), based on the content of the units on which they proposed to work during the week. There were 14 participants in the Earth/Space Science group, 15 participants in the Physical Science group, and 16 participants in the Life Science group. Each mega-strand group divided into two or three *break-out* groups, and most of the problem-based learning activities occurred in these break-out groups; each group typically studied one problem, called a *content dilemma*, each day (we use the terms *content dilemma*, *problem*, and *PBL materials* interchangeably). In order to guide the process of problem-based learning in each group, a two-sided instructional card was given to all facilitators and participants; this card explained the PBL process (see Appendix A). The daily schedule allotted time for participants to both work with content dilemmas through the PBL process and develop their instructional lessons or units. The two activities were not necessarily connected.

To begin the PBL process, facilitators provided each participant in the group the first page of the problem; this page introduced the problem and gave a small amount of information. At that point, facilitators asked group members to list facts about the problem, and one of the facilitators recorded their responses on chart paper. After a time, facilitators asked the group to list hypotheses and then learning issues (questions) about the problem in the same manner. Next, facilitators distributed page two of the dilemma to the participants; page two revealed more information about the problem, and the process of listing facts, hypotheses, and learning issues was repeated. Facilitators then gave page three to the participants; this page listed resources for the participants to use in solving the problem, and they could supplement that list with other sources if they desired. The distribution of page three began the self-directed learning phase, and one or more facilitators guided this process. Some self-directed learning phases consisted of short periods of time in which participants only made use of the resources listed on page three. At other times, participants decided to break into small groups to build models or conduct Internet searches. Occasionally, participants decided to stay together as a group to build models together. After the self-directed learning time, participants gathered back together to report on their learning. At that time, they listed relevant big ideas of science content and related curriculum objectives. After doing this, facilitators passed out the fourth page, which contained the state and/or national standards associated with the content that the facilitators had hoped participants would explore with the problem.

After the end of the *PWC*, twenty-three of these teachers completed a second week of PD, called the PBL *Focus on Practice* (this began the second phase of the PD). The two primary activities during this second week consisted of more time for teachers to develop their own instruction around their focal area (in response to what they learned during the first week) and more PBL sessions. The PBL exercises in the second week emphasized issues of teaching (e.g. assessment, activity structures, student understanding in the context of science) more than content knowledge.

The teachers that participated in the *Focus on Practice* PD continued to meet each month in small groups throughout the 2005 – 2006 academic year; these groups used the PBL process to investigate their own self-identified problems of practice. During these meetings the teachers receive support from other teachers, project staff, and researchers, and the teachers will present their findings at a project-wide year-end poster session in May of 2006.

The Problems

A total of 25 problems were developed for use in the *Professional Working Conference*: seven for the Earth/Space Science group, ten for the Physical Science group, and eight for the Life Science group. For this analysis, we selected the content dilemmas that participants rated the highest and lowest in one break-out group from each of the Life Science and Earth/Space Science mega-strands, as we report below. We selected these break-out groups because we had recorded their content dilemma sessions on videotape, which provided us with the advantage of more data for our analysis. As well, the facilitators were the same for each problem in these two break-out groups, which we anticipated could reduce the number of variables in our analysis.

As we alluded to previously, each content dilemma consisted of four pages. Each page revealed more information to the participants in an effort to scaffold their learning. For this paper, we analyzed four of these content dilemmas; we include these as Appendices B-E.

Participants in the Cells break-out group of the Life Science mega-strand rated the *Hamm Brothers* and the *Blood Doping* dilemmas lowest and highest, respectively. In the *Hamm Brothers* dilemma, the facilitators presented the participants with two editorials about whether two twins on the US Olympic Men's Gymnastics Team were identical or fraternal. The problem directed participants to study cell reproduction and the various processes that produce twins. In the *Blood Doping* dilemma, the facilitators presented the participants with a scenario in which two students in a classroom were arguing about whether Tyler Hamilton should have been excluded from bicycle racing competitions based on the results of a controversial test for a performance-enhancing medical procedure. The problem directed participants to discuss ways to use the students' interest in the issue to motivate learning in science class.

Participants in the Astronomy break-out group of the Earth/Space Science mega-strand rated the *Length of Day* and *Moon Phases* problems lowest and highest, respectively. In the *Length of Day* problem, participants were given the opportunity to examine a children's book about Fairbanks, Alaska. Then, participants were given temperature and sunrise/sunset data on Fairbanks and Lansing, Michigan. The problem directed participants to explain the reason for the vast differences in light and temperature between the two cities. The *Moon Phases* problem required participants to develop an explanation for the reason moon phases occur. In both astronomy problems, participants chose to make models to aid them in explaining their solutions.

Data Collection and Analysis

We obtained participant ratings of the dilemmas from a survey we developed to measure teachers' perceptions of the content dilemmas. In that instrument, we asked participants to judge the usefulness of the content dilemmas as "Not helpful," "Somewhat helpful," or "Very helpful," and to write comments about their ratings. We assigned two points for each rating of "Very," one point for each rating of "Some," and zero points for each rating of "Not," and we then calculated a mean score for each problem. In Table 2, we list the scores for the four dilemmas we will study. <Insert Table 2 about here.> We examined the ratings for all 25 dilemmas to see whether trends existed across the ratings, such as whether dilemmas presented later in the week were rated more highly than those presented earlier, or whether elementary teachers and secondary teachers rated dilemmas differently. We found no such correlations.

After we synthesized the criteria from the literature (see Table 1 above), four members of the research team obtained text copies of the four problems that participants received during the PD event. We evaluated each of the content dilemmas according to the first five criteria. We chose not to evaluate the content dilemmas in terms of "engagement" at that time, because we judged this characteristic to not be apparent in the written dilemmas. Instead, we waited and looked for evidence of engagement while analyzing the video. Each researcher scored each problem according to each of the first five criteria (authentic and relevant, appropriate in regards to curriculum payoff, controversial, appropriately complex, and clear about participants' roles), using a scale of 0 to 2. For example, we scored problems regarding their authenticity with a 0 (*not at all authentic*), 1 (*somewhat authentic*) or 2 (*authentic*). After scoring the problems, the four researchers met to share their scores and to justify them according to the literature. As a part of this process, the researchers revised and expanded the first five criteria for evaluating the content dilemmas, and we came to consensus on how to rate the dilemmas for each of the criteria. We list our ratings of the content dilemmas in Table 3. <Insert Table 3 about here.>

We also collected each of the break-out group's notes that facilitators had scribed on chart paper, and we used these in our analysis. The notes included facts, hypotheses, learning issues, recommended actions, and big ideas of the science content pertaining to the content dilemmas. We analyzed the notes according to the number of facts, hypotheses, and learning issues that participants generated, and we made observations about whether the facts, hypotheses and learning issues allowed the participants to engage in rich discussions or rigorous self-study about the science content.

In addition, we recorded group discussions of the content dilemmas on videotape. Two researchers analyzed each video. The researchers used a template to prompt them to look for trends in the facilitation and the engagement of participants in the problems, as well as examples of participants engaging in the goals and outcomes that the PBL process intends to generate. Specifically, the researchers sought to make comparisons between the problems in terms of whether and how each problem lent itself to the goals and outcomes of the PBL process. As well, researchers looked at the video for evidence of participants' engagement.

Results

Comparison of Two Content Dilemmas Based on Criteria from the Literature

After reaching consensus about how to score each of the problems according to the criteria we had defined from the literature on a scale of 0-2, each of the problems were rated as seen previously in Table 3.

Cell biology problems. When we applied our criteria for what constitutes a good problem to the *Hamm Brothers* dilemma, we concluded that this dilemma lacks many of the characteristics of strong problems. The dilemma lacks authenticity and relevance in that it precludes participants from taking an active role in its solution (because they lack the ability to conduct a DNA test on the Hamm twins or obtain the results of one). The dilemma is controversial in the sense that the outcome of the inquiry is unknown; however, the controversy is not of substantial interest to the field of science, nor does it really allow participants to develop rival, verifiable hypotheses. This inability to develop competing hypotheses indicates evidence of a lack of complexity in this dilemma. Further, because participants are not able to be involved in the problem-solution, participants could be unsure about their roles in the dilemma, which may interfere with participants' motivation to make any conclusive decisions about the dilemma. With respect to curricular appropriateness, the *Hamm Brothers* dilemma does not effectively encourage participants to entertain high levels of thinking about cell reproduction; further, its attempted connections to classroom benchmarks are thin. We conclude, based on our analysis of the written problem, that the *Hamm Brothers* dilemma lacks authenticity, controversy, and complexity, and it fails to be pedagogically effective or clear about participants' roles.

In contrast, we found that the *Blood Doping* dilemma more closely aligned with our criteria for a good problem. Setting the dilemma within a classroom provided an authentic and relevant context for our teacher-participants. The problem required participants to consider biology content within a given classroom situation; we anticipated this could make participants' roles easily imaginable. With respect to the science content of the dilemma, the controversy around the practice of blood doping actually exists, and it has the potential to allow participants to pursue legitimate inquiry due to the complexity of the issue. We judged the dilemma to be pedagogically effective because of the multiple, viable hypotheses that participants could generate, and because the decisions arrived at by participants could be supported in scientific literature. Further, the dilemma naturally allows for a discussion of important scientific issues, such as the function of red blood cells and the processes of circulation and respiration, as well as ethics, sportsmanship, and healthy living. We concluded that the *Blood Doping* problem provided authenticity and relevance, curricular appropriateness, controversy, and clarity of participants' roles.

The *Blood Doping* dilemma did not receive the highest rating in complexity, however. It had the potential to be appropriately complex, but the participants were given too much information, and therefore complexity was compromised. The participants could have had a greater opportunity for self-directed research had they been required to find their own resources or to choose among several resources to use in solving the problem, instead of being given two, very complete resources.

Astronomy problems. The *Length of Day* problem had potential for addressing big ideas related to the change in seasons. Yet, the curriculum payoff was limited by the fact that it was not written to address some of the goals of problem-based learning. Although participants could work together to learn the content, we did not anticipate that it would require them to discuss and debate competing hypotheses. The dilemma could be less relevant to educators who do not teach this content to their students. As written, it did not appear to address authentic issues of pedagogy. This problem is not current; scientists do not presently debate the reason for these changes. These facts limited the complexity and authenticity of the problem.

When applying the criteria to the *Length of Day* dilemma, we decided that this problem was not controversial or clear about participants' roles. Since this problem has one correct solution, it seemed unlikely to be controversial. There was not potential for competing, viable hypotheses about which teachers would be likely to disagree. Although participants were asked to answer the question about what causes the different change in temperatures at these two locations and report their answer to the whole group, the problem was not written to include the participants as stakeholders in the solution to any problem. They were presented with a question that had one scientifically accepted answer and were not asked to think about the implications of the content for classroom teaching. We determined that the *Length of Day* dilemma lacked controversy and failed to be clear about participants' roles, yet was somewhat authentic and relevant, pedagogically effective, and complex.

Each criteria rating for the *Moon Phases* dilemma was an exact match for the *Length of Day* problem. Like the *Length of Day* problem, we decided that the authenticity of the *Moon Phases* dilemma depended on whether or not participants currently teach students about what causes the changes in the moon's shape. If this content was addressed by one's curriculum benchmarks, then the problem is likely to be more relevant to that person. Content learning regarding how the position of the sun, earth, and moon affect the phase change likely will be addressed by this problem. Yet, collaboration and development of participants' problem-solving abilities is only likely to occur if participants do not already know the answer to this problem. Also, the suggestion that the position in the sky has something to do with this phenomenon is fairly leading and could limit these outcomes of the PBL process. This problem is not controversial because there is only one right, scientifically accepted answer to this question. Equally viable hypotheses are unlikely to be debated by the participants. In addition, participants are not given a role to fill as they determine the answer to what causes the changes in the moon phases. We concluded that the *Moon Phases* dilemma was somewhat authentic, pedagogically effective, and complex, but that it lacked controversy and was not clear about participants' roles.

Facilitator/Board Notes from Discussions of the Content Dilemmas

Differences existed among the notes recorded by the facilitators in the various dilemmas. For example, participants in the *Hamm Brothers* dilemma listed only one hypothesis, "They are identical twins". The same participants listed three hypotheses while working through the *Blood Doping* dilemma. Those hypotheses were, "He was blood doping"; “A medication caused the false positive"; and "They used a more controversial test". The difference in number of viable hypotheses indicates that participants found the *Hamm Brothers* dilemma to stimulate less controversy than *Blood Doping* dilemma. Participants did not generate any hypotheses or learning issues regarding curriculum or pedagogy, even though the *Blood Doping* dilemma opened up this possibility to them. However, we surmise that this is because the dilemmas presented to the participants prior to this one did not allow the option to consider pedagogy, and we believe participants had already formed habits regarding the PBL process (which did not include thinking in terms of classroom applications).

Differences also appeared in the number and quality of learning issues generated by those same two problems. While discussing the *Hamm Brothers* dilemma, participants generated eight learning issues. Many of those were overly simplistic, were not necessary for learning the content that the dilemma was intended to cover, or were not answerable with the available resources. For example, a facilitator wrote the question, "Same DNA?" on the chart paper when a participant asked whether the Hamm twins had the same DNA. Participants did not possess the means to find the answer to that question. In comparison, participants generated fifteen learning issues while discussing the *Blood Doping* dilemma, and these issues were researchable and pertained to relevant science content. For example, participants asked, "What is blood doping?"; "What could cause false positives?"; "When is the best time to test?"; "Is it harmful to the individual?"; and "Is it beneficial beyond use for end stage kidney disease?". All of these learning issues could lead to self-directed study on red blood cells and body systems, within the context of an idea that was new to the participants.

While the number of hypotheses generated by the participants who worked on the two Earth/Space Science dilemmas was similar, they were quite different in quality. While working through the *Length of Day* dilemma, participants generated six hypotheses. We believe these particular hypotheses may indicate that at least some of the participants already had knowledge of the solution to the problem. For example, some of the participants offered: "The amount of daylight causes the change of temperature"; "The tilt of Earth impacts the amount of daylight"; and "The closer to Arctic Circle the more extreme the change in temp and daylight”. However, a minority of participants offered hypotheses to the group, and it is possible that they were not representative of the group as a whole. Another possibility is that those hypotheses naturally stemmed from the data presented to the participants, which showed changes in daylight and temperature in Fairbanks and Lansing over time.

While working on the *Moon Phases* dilemma, participants produced seven hypotheses. One was that "The E's moving causes the moon to appear to be getting smaller/larger." However, participants offered several of the hypotheses in terms of common (student) misconceptions. For example, "We see the moon because of light from E"; "Clouds blocking the light cause the phase changes"; and "Moon is only out at night because the sun goes down, making the moon come up" were all mentioned along with a discussion of ideas students might have in these teachers' classrooms.

The fourteen participants working on the *Length of Day* dilemma generated nine learning issues. Some of these were more relevant to the solution of the dilemma, and some were less relevant. For example, "Tilt of Earth's axis each month"; "Angle of sun's rays for each month"; and "How does tilt of Earth affect daylight" could lead participants closer to the answer. However, "Is data consistent for each year" and "Who did the measuring of sunrise set etc.?" may provide additional information, but would not directly lead participants to the solution to the problem.

The eight participants working on the *Moon Phases* dilemma suggested seven learning issues. They seemed to start by thinking about the content in terms of what students would do with it, asking, "What's producing light?" and "Why change shape when solid sphere?". However, as the discussion progressed, they seemed to be asking deeper questions regarding their own content knowledge, such as "Who's moving?"; "Why moving?"; "What path does it follow?"; and "Where does it revolve?".

The differences we observed in facilitator board notes seem to align with our ratings of the dilemmas according to the criteria we synthesized from the literature. The two astronomy problems received identical ratings according to our established criteria, and participants generated similar numbers of hypotheses and learning issues in the enactment of the problem. In contrast, the cell biology problems, which received very different ratings when compared with the criteria, produced different numbers of hypotheses and learning issues. In general, these results align with participant ratings in the area of astronomy, as expected.

Participants' Survey Comments

In addition to ratings, we asked participants to provide comments on the survey about the dilemmas and how they rated them. For the *Hamm Brothers* dilemma, four participants wrote comments on the survey. Participants wrote, "Very confusing; not at all interesting"; "Wasn't rich at all"; "Bored me. It was not something I cared much about"; and "Writing all the facts makes this take forever. Interesting questions about clones though". These comments speak to the authenticity and relevance of the dilemma, as well as its complexity and engagement. We interpret these comments to mean that participants were not invested in the problem. In addition, these comments seem to indicate that the text of the problem provided too much information and did not require participants to gather enough of their own information.

Most of the comments written by the participants regarding the *Blood Doping* dilemma were more positive in nature. Participants wrote, "I really enjoyed the discussion"; "Lots of good questions came up and it was engaging—plus I like biking!"; and "Rich concepts and no right answers. Challenging. Something to think about ethically too". These comments speak to the relevance, complexity, curriculum payoff, and controversy in the *Blood Doping* dilemma. A fourth participant wrote, "Way to [*sic*] difficult for 7th graders," indicating that he or she would have found a dilemma that focused on content relevant to his or her particular grade level more engaging.

The importance of the authenticity and relevance of the dilemmas emerged again when participants responded to general questions on the survey regarding how facilitators could improve the dilemmas and the conference as a whole. Several teachers in the Life Sciences mega-strand said that the dilemmas and the conference would have been improved if the content of the dilemmas was more closely aligned with the science content of the units they were developing during the course of the PD.

Two participants from the Earth/Space Science mega-strand who worked on the *Length of Day* dilemma commented on learning new facts about the movements of the earth and sun. Participants made other positive comments regarding the opportunity to work in teams and the connection the dilemma made to children's literature. Two participants, however, found the content to be "confusing", and one pointed to the children's book as a possible source of confusion. Another claimed to have already known the science content of the dilemma, but stated, "(I) liked the activities we did. I could use them in my classroom."

Participants who worked on the *Moon Phases* dilemma all had very positive comments about their experience with it. Five participants specifically stated that they felt their knowledge of the reason for the moon phases had increased due to their participation in the dilemma. This speaks to the relevance, complexity, and curriculum payoff of the dilemma. Two stated that they felt like the models they used for finding answers to their learning issues were ones that they could use in the classroom. One participant said, “They were so very relevant."

When teachers in the Earth/Space Science groups were asked how the dilemmas and workshop could be improved, a few stated that they would have liked dilemmas that were more explicitly connected to the science content that they taught at their grade level. We believe this speaks to the need for the content in the dilemmas to be obviously applicable to teachers in the particular classrooms in which they currently teach.

In conclusion, participants' comments on the survey provide some evidence that these criteria are important in terms of participants' satisfaction. Participants indicated that authenticity and relevance, controversy, complexity, and clarity about their roles are important to a problem that they perceive as helpful.

Video

The purpose of analyzing the video was to see if participants were engaged in the process of problem-based learning, namely, generating facts, hypotheses, and learning issues, engaging in self-directed study, collaborating with peers to solve problems, making a decision regarding clinical reasoning in the classroom, and backing up their solution with evidence. While we are aware that other issues, such as facilitation, influence these, we looked for evidence that the problem did or did not lead naturally to these PBL goals. In short, we attempted to study how the problems themselves influenced the PBL process.

Hamm Brothers. The video of participants working on the *Hamm Brothers* dilemma was characterized by confusion, lack of participants' roles, and indifference. The following excerpt is illustrative of the type of confusion participants experienced when working on the *Hamm Brothers* dilemma:

Jeff: What's the question we're looking at here? Is it are they fraternal?

Fac-1: There are a couple of questions here. One is what is the difference between identical and fraternal.

Fac-2: That's a learning issue

Fac-1: Yeah, and does it make any difference? It's something to think about, but.

Jeff: When I read that I'm trying to figure out what's the, what are we trying to figure out? So what's the problem we're looking at?

Fac-1: What do you guys think? Would be, does that sound reasonable?

Jeff: So, well, that makes the hypothesis easy, because either they are identical or they're not. So we have two hypotheses.

Ann: Wait a minute. The article says that they're not.

Mary: They're not identical.

Ann: It says, "The Hamm twins and their father responded that they are not identical twins but fraternal twins." So we know what kind of twin they are.

From this excerpt, it is apparent that the participants are unclear of their roles and of the problem. As well, the statement, "That makes the hypothesis easy," is evidence of the lack of complexity and controversy in the problem. Ann points out that, according to the family in question, there is no problem, and therefore nothing to solve. Because of this lack of complexity, participants who worked on the *Hamm Brothers* dilemma did not need to do any self-directed learning. While a couple of learning issues were generated, participants chose not to do any self-directed research to answer them, indicating a lack of engagement in the content of those questions. As well, the unclear roles and lack of complexity led to a lack of collaboration. This is also evidence that the dilemma, which we did not rate highly in terms of the criteria for a good problem, was not engaging and did not contribute to the goals of the PBL process or the science content curriculum.

Blood Doping. This dilemma was the first one that placed the participants in the role of a teacher and addressed pedagogical decision-making. Participants focused on the science content and talked minimally about using the controversy in their own teaching. For the self-directed learning portion of this dilemma, the participants were provided with two articles from the Internet as resources. Since more information was provided than they would normally have when teaching, there was no need for participants to engage in rigorous academic work in which they sought out their own sources to answer the identified learning issues. As a result, the complexity, authenticity and relevance written into the problem were lessened in its enactment.

During this session, one of the facilitators tried to bring the conversation back to the pedagogical issue:

Fac-2: So how would you envision using this with your students if your students were posed with this problem, and they wanted to know, well, why is it, you know, you just asked a really good question. It's the kind of thing students would ask too. How would you go about answering it? Would you tell them outright?

Alex: Saying, ok, here's the scenario. You tell me what the benefit of this would be.

Julie: Make it an assessment?

Alex: Yeah. So that I can see, can they take this, you know, I mean, big deal for them. They learn one day, oh, red blood cells carry oxygen. Great. Now can you apply it to something, how this is actually used?

Again, after the participants had completed the dilemma, a facilitator asked them to discuss the use of this or other dilemmas in their own classrooms. This part showed the highest level of engagement during this dilemma. The participants discussed pedagogy, but were not required to support their reasoning with evidence. Because of that, the full potential for this problem, as written, was not realized as it was enacted. Regarding the written problem, we conclude that the *Blood Doping* dilemma contained too much information and therefore lacked complexity, which interfered with the PBL process. However, because the *Blood Doping* problem was authentic and relevant, controversial, appropriate in regards to curriculum payoff, and clear about participants' roles, it was more engaging than the *Hamm Brothers* dilemma and lent itself more readily to the PBL process.

Length of Day. Analysis of this video shows that the PBL process was not complete. Participants generated facts, hypotheses and learning issues, and they engaged in collaborative self-directed study. However, the later stages of the PBL process regarding making a decision and backing up the solution with evidence were absent from the process.

Another difficulty arose because, although the problem was written to include both an explanation of both the length of daylight and of temperature, when working through the self-directed learning phase, participants became confused about what they were being asked to explain. As a part of the introduction to the problem, facilitators suggested that participants build models, so that is what participants began to do. However, they were confused about what the models were for. For example, see this excerpt from the self-directed learning phase:

Fac-3: So this is what I'm going to suggest that you refocus on right now and that is trying to find the answer to that first question. And what you need to do because any one of the four of you may be asked to give an explanation of your model. Is make sure you understand about…

Bess: Answer to the first question. About the tilt.

Fac-3: About the temperature.

Bess: Oh, No, We started on light.

Jill: I thought we had the light figured out.

Fac-3: So you need to think about light in Fairbanks, light in Lansing, and make a model that shows (can't hear) about

Bess: This isn't good enough for a model?

Fac-3: It could be.

After participants built their models, they shared them with the whole group. However, the problem was concluded without the participants having to make a decision regarding clinical reasoning, or having to back up their findings with evidence. Therefore, we conclude that the video shows that the problem did not lend itself to the PBL process due to a lack of clarity of participants' roles and lack of controversy.

Moon Phases. This problem, as written, included some of the criteria for a good dilemma. Based on our criteria, we rated the *Moon Phases* problem the same as the *Length of Day* problem, and so did the participants who were in both the *Moon Phases* and the *Length of Day* problem discussion groups. However, we argue that the facilitation made the *Moon Phases* problem more authentic and relevant to the participants than we predicted by allowing for discussions that involved instructional decision-making.

The video shows that many parts of the PBL process were included: generating facts, hypothesis, learning issues, engaging in self-directed study, and collaborating with peers to solve problems. Although the authors did not write issues of pedagogy into the dilemma, participants continually brought them up. The facilitator built on their ideas and included connections to teaching strategies and student thinking, making the experience of the problem relevant and authentic to the participants. For example, the facilitator asked questions such as, "Do you have some hypotheses that children would bring to this?", and she was encouraging of teachers when they brought up issues of classroom teaching:

Linda: I at first thought it would be good to read the book first, but now, I think you could hand the book out during this question phase. Like, let's find some facts about this.

Kevin: They do like to do that research. They think of themselves as researchers.

Sherry: Actually I was thinking of a different approach, just kind of polling them to see what they know as facts first. Then perhaps read the book. I want to get a picture of their thinking first.

We conclude that the PBL process could have been more fully realized if the problem had been more controversial. However, we see differences in the way the problem was written and the way it was facilitated. Participants continually brought up classroom issues, and the facilitator encouraged this, which lends support to the idea that problems need to be authentic and relevant. As well, the facilitator brought clarity to the participants' roles, even though it was not written into the problem, by encouraging them to place themselves in the role of teachers. This lends support to the idea that problems need to be authentic and relevant, as well as clear about participants' roles.

Discussion

We noted at the beginning of this paper that this study was exploratory in nature. Consequently, we offer our recommendations for the writing of problems for use in science teacher PD in light of the preliminary nature of the study. Nonetheless, we believe that our synthesis of the literature, coupled with the supporting data we collected from the conference participants and sessions, provides us with some ground on which to make some suggestions regarding the writing of problems. In this section, we offer our recommendations in terms of amending two of the problems that we analyzed in this paper. We believe our amendments would enhance the potential of those problems to realize more of the goals of the PBL process. We conclude this section by pointing to the need for further research and summarizing the arguments that we made in this paper.

Revisiting the Content Dilemmas

The existing literature on writing problems, as well as our analysis of the data that we collected from the conference participants and sessions, suggest that problems need to be authentic and relevant to the participants, controversial, complex, and clear about participants' roles. In the context of science teacher professional development, we interpret those findings to mean that the content dilemmas (problems) need to incorporate instructional decision-making opportunities. In other words, we recommend that content dilemmas not only engage participants in exploring content that is relevant to classroom science teachers, but also that they include opportunities to think about the content in terms of classroom instructional practices. We think that placing the setting for the content dilemmas in the context of schools or classrooms, or writing into the dilemmas the opportunity for making instructional decisions, will enhance the potential of the dilemmas to realize all of the components of the PBL process. We argue that such settings and decision-making tasks could make the problems ones that teachers would be more likely to experience or be familiar with and it would be more likely that the problem would be within the scope of their powers to solve. We anticipate that such problems would enhance the motivation of participants to engage in problem-solving and professional decision-making tasks.

Length of Day. While working through the *Length of Day* dilemma, participants built three-dimensional models to explain the reason for the changes in the temperature and the amount of daylight in Alaska and Michigan. This is important content knowledge for teachers to possess, and many teachers did not begin the conference with deep knowledge of this content. However, we propose that this problem would be more effective if it delivered the same content but was set within a school or involved instructional decision-making tasks. For instance, a component of the problem could require teachers to select an appropriate model of the sun/earth/moon system for teaching their students about the change in seasons. This amended version of the dilemma could possibly afford a deeper investigation of the content, as it could lead the participants to engage in a debate regarding the benefits and detriments of a variety of model types. Further, participants could discuss how each model shows or does not show the causes of the change in seasons, what student misconceptions each model might create or debunk, and what important aspects of the content each model clarifies or highlights. In addition, the amended version could motivate participants to make a decision and defend their reasoning with evidence, which would improve the dilemma in terms of its pedagogical effectiveness.

Blood Doping. According to the literature, a problem is appropriately complex if it requires participants to collaborate and engage in inquiry in efforts to solve it. In addition, it needs to provide no more or no less information than the participants would have if they actually encountered the problem. Based on the text of the problem, we anticipated that the *Blood Doping* dilemma would be pedagogically effective. However, although participants engaged in the process of identifying facts, hypotheses, and learning issues, we found little evidence of peer collaboration or problem solving on the video. We interpreted this to be a result of the materials of the dilemma providing the participants with too much information, and thereby not encouraging them to collaborate or engage in problem solving to search for solutions. Therefore, we would amend the *Blood Doping* dilemma by reducing the amount of information provided to the participants about the process of blood doping. We anticipate that this amendment would require participants to locate and evaluate materials on their own and with their peers. We believe this would enhance the opportunities for the participants to engage more fully in the PBL process and realize its intended outcomes.

The Need for Further Research

We have recommended that problems written for use with science teachers in professional development contexts be set in classrooms or schools and allow participants an opportunity to engage in professional decision-making tasks. We hypothesize that these types of dilemmas, ones that introduce issues of pedagogy, hold more potential for the realization of all components of the PBL process than do dilemmas that include only issues of content and ignore pedagogy. We base this hypothesis on our own considerations of the literature and on the data we collected and analyzed from conference participants and sessions. However, we believe that there is a need for further research regarding that hypothesis. Researchers need to conduct studies that gather empirical evidence to compare the effectiveness of problems that address only content and no pedagogy to ones that address only pedagogy and little deep content. Researchers need to explore the balance of content and pedagogy that would most likely be engaging, and at the same time, result in the acquisition of deep content knowledge by participants.

There also exists a need for researchers to study whether all of these criteria are equally important and necessary in the context of science teacher professional development. In addition, there is a need for more research concerning how well a framework for the analysis of written problems correlates with measures of participants' changes in content knowledge and clinical reasoning skills. We have begun this exploration, but there is a need for more of this type of work.

Summary

In this paper, we wrote that a literature exists regarding the characteristics of good problems, but we did not find any research that reported empirical work, and we found little research that applies to the context of the professional development of teachers. We synthesized the work on the writing of PBL problems (for many contexts) and cases (in educational contexts) in order to create a framework for evaluating PBL materials. We evaluated four content dilemmas (PBL problems) in light of our framework, and we drew conclusions about the potential effectiveness of those dilemmas. We analyzed participant surveys, facilitator board notes, and discussions (captured on videotape) associated with those content dilemmas in an effort to see if those data corroborated our evaluations of the problem texts. We argued that the data that we obtained from the conference proceedings supported our conclusions regarding the effectiveness of the problems. We concluded by considering the implications of our findings on the writing of PBL problems in the context of science teacher professional development.

References

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine, 68*(1), 52-81.

Allen, D. E., Duch, B. J., Groh, S. E., Watson, G. B., & H. B. White, I. (2003). Scaling Up Research-Based Education for Undergraduates: Problem-Based Learning. In *Reinvigorating the Undergraduate Experience: Successful Models Supported by NSF's AIRE/RAIRE Program.* Retrieved Retrieved 9/1/2004, 2004, from http://www.cur.org/publications/AIRE_RAIRE/printer/delaware.asp.

Astronomical Society of the Pacific (ASP) (2006). The *Astronomical Society of the Pacific Homepage*. Accessed on March 15, 2006 at http://www.astrosociety.com.

Barnett, C., Goldstein, D., & Jackson, B. (1994). *Fractions, decimals, ratios, and percents: hard to teach and hard to learn*? Portsmouth, NH: Heineman.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*. New York: Springer.

B.B.C. (2004). *Hamilton fails dope tests: Olympic time trial champion Tyler Hamilton faces a ban for doping after failing blood tests in Athens and last week's Tour of Spain*. Retrieved on May 20, 2005 from http://news.bbc.co.uk/sport1/hi/other_sports/cycling/3675726.stm.

Beckham, D. (n.d.) *Blood doping: Is it really worth it?* Retrieved on May 20, 2005 from http://www.tc.cc.tx.us/~mstorey/beckham.html.

Butler, S. M. (2003). Designing a technology-based science lesson: Student teachers grapple with an authentic problem of practice. *Journal of Technology and Teacher Education, 11*(4), 463-481.

Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction, 22*(1), 61-79.

Carter, K. (1999). What is a case? What is not a case? In M. A. Lundeberg, B. B. Levin, & H. L. Harrington (Eds.), Who learns what from cases and how? The research base for teaching and learning with cases (pp. 165–175). Mahwah, NJ: Lawrence Erlbaum Associates.

Chernobilsky, E., DaCosta, M. C., & Hmelo-Silver, C. E. (2004). Learning to talk the educational psychology talk through a problem-based course. *Instructional Science, 32*(4), 319-356.

Delisle, R. (1997). *How to use problem-based learning in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Derry, S. J., Seymour, J., Fassnancht, C., & Feltovich, P. (2001). *Tutoring and knowledge construction during problem-based learning: An interaction analysis.* Paper presented at the 2001 Annual Meeting of the National Association for Research in Science Teaching, St. Louis, MO.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction, 13*(5), 533-568.

Duch, B. J. (2001). Writing problems for deeper understanding. In B. J. Duch, S. E. Groh & D. E. Allen (Eds.), *The power of problem-based learning* (pp. 47-58). Sterling, VA: Stylus Publishing.

Gordon, R. R., Rogers, A. M., Comfort, M., Gavula, N., & McGee, B. P. (1991). A taste of problem-based learning increases achievement of urban minority middle-school students. *Educational Horizons, 79*(4), 171-175.

Greenwood, G. E., & Fillmer, H. T. (Eds.). (1999). *Educational psychology cases for teacher decision-making*. Upper Saddle River, NJ: Prentice Hall, Inc.

Harrington, H. L. (1995). Fostering Reasoned Decisions: Case-Based Pedagogy and the Professional Development of Teachers. *Teaching & Teacher Education, 11*(3), 203-214.

Hansen, A. J. (1997). Writing cases for teaching: Observations of a practitioner. *Phi Delta Kappan, 78*, 398-403.

Herreid, C. F. (1994). Case studies in science: A novel method in science education. *Journal of College Science Teaching, 23*(4), 221-229.

Herreid, C. F. (1997). What makes a good case? Some basic rules of good storytelling help teachers generate student excitement in the classroom. *Journal of College Science Teaching, 27*(2), 163-165.

Hmelo-Silver, C. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, (16)*3, 235-266.

Journey North (2005). *The Mystery Class* website. Accessed online on March 15, 2006 at http://www.learner.org/jnorth/mclass/spring2005/.

Kelson, A. C. (2004, November). *The power of problem oriented learning.* Paper presented at the Problem-Based Learning Professional Development Project, East Lansing, MI.

Levin, B. B. E. (2001). *Energizing teacher education and professional development with problem-based learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Levin, J. (2004a). *The 2004 Olympics: Are Paul and Morgan Hamm identical or fraternal twins?* Posted on August 18, 2004 on http://www.slate.com/id/2105357/.

Levin, J. (2004b). *The 2004 Olympics: Are Paul and Morgan Hamm identical twins? Their parents say no. Science says yes.* Posted on August 19, 2004 on http://www.slate.com/id/2105372/.

Levin, J. (2004c). *The 2004 Olympics: Why the U.S. men's gymnastics team is un-American*. Posted on August 17, 2004 on http://www.slate.com/id/2105271/.

Lundeberg, M. A., Levin, B. B., & Harrington, H. L. (Eds.). (1999). *Who learns what from cases and how: The research base for teaching and learning with cases*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lundeberg, M. A., Mogen, K., Bergland, M., Klyczek, K., Johnson, D., & MacDonald, E. (2002). Case it or else! Fostering ethical awareness about human genetics. *Journal of College Science Teaching, 32*(1), 64-69.

Michigan Department of Education (MDE) (2000). *Michigan curriculum framework science benchmarks*, retrieved online on March 15, 2005 from http://www.michigan.gov/mde/0,1607,7-140-28753_38684_28760---,00.html.

Newman, M. (2004). *Problem Based Learning: An exploration of the method and evaluation of its effectiveness in a continuing nursing education programme.* Middlesex University.

Newman, M., Ambrose, K., Corner, T., Vernon, L., Quinn, S., Wallis, S., et al. (2001). *The project on the effectiveness of problem-based learning (PEPBL): A field trial in continuing professional education.* Paper presented at the Third International Inter-disciplinary Evidence-Based Policies and Indicator Systems Conference.

Shulman, J. H. (1996). Tender feelings, hidden thoughts: Confronting bias, innocence and racism through case discussions. In J. Colbert, P. Desberg, & K. Trimble (Eds.), *The* *case for education: Contemporary approaches for using case methods*, pp. 137– 158. Boston: Allyn & Bacon.

Sullivan, J. A. (1999). Another drug war. Retrived on May 20, 2005 from http://whyfiles.org/090doping_sport/3.html.

Torp, L., & Sage, S. (2002). *Problems as possibilities: Problem-based learning for K-16 education*. Alexandria, VA: Association for Supervision and Curriculum Development.

U.S. Naval Observatory (USNO) (2006). *Complete sun and moon data for one day*. Accessed online on March 15, 2006 at http://aa.usno.navy.mil/data/docs/RS_OneDay.html.

Weather Underground, Inc. (2006). *Wunderground.com*. Accessed online on March 15, 2006 at http://www.wunderground.com.

Appendix A

PROBLEM-BASED LEARNING
CONTENT DILEMMA ANALYSIS PROCESS

BEGINNING

Introductions
Climate setting
- Open thinking/everyone contributes
- Silence is assent.
- Everyone speaks to the group/No side conversations.
- Facilitator will ask the group questions in order to clarify and keep the process going.
- A scribe will record the group's thinking.

STARTING A NEW PROBLEM

- A volunteer reads the first page of the content dilemma aloud.
- Participants identify facts/information, hypotheses, questions/learning issues.
- The reader reads the second page of the content dilemma aloud.
- Participants add to and modify facts/information, hypotheses, questions/learning issues.
- Participants prioritize learning issues, identify possible resources, and develop an action plan for independent learning.

Information/facts	Hypotheses	Questions/ learning issues
• From content dilemma & research • What do we know?	• Tentative explanations based on information/facts • What do we think is going on?	• What needs to be learned in order to distinguish btw or generate hypotheses • What do we need to know?

SELF-DIRECTED STUDY

Participants engage in independent learning.

RETURN TO CONTENT DILEMMA

- Participants add new learning to information/facts.
- Participants apply new learning to hypotheses and questions/learning issues.

FOLLOW UP

- Participants summarize the resolution of the content dilemma and review content lists.
- Participants identify how new learning integrates with previously learned concepts and with unit content.

DEBRIEFING

- Reflection on the reasoning process.
- Reflection on the group process and facilitation.

GOALS FOR PBL PROFESSIONAL DEVELOPMENT
Premise: Teachers are professional clinicians and their development must mirror training of the clinical professional. *"…Assess, diagnose, prescribe and adjust practice to reflect new research, training and experience – that's what a modern clinical professional does. The job description not only fits physicians who see patients in clinics, it precisely defines the work of teachers who see students in classrooms. One conceptual answer {to improving education} is for society to treat teaching like the modern clinical profession that the nation needs it to be."* (Hinds, M. (2002) Teaching as clinical profession: a new challenge for education. New York: Carnegie corporation of New York.) Clinical reasoning for teachers • A systematic approach to solving real-life problems • An extensive, integrated knowledge base that can be recalled and flexibly applied to other situations • Effective self-directed learning skills • Attitudes and skills for effective team work • The life long habit of approaching a problem with initiative and diligence and drive • Habits of self-reflection and self-evaluation (Southern Illinois University School of Medicine, Springfield, IL)
Goals: • Promotion of clinical reasoning of teachers. • Curricular revision and implementation that includes deepening teachers' subject matter knowledge and pedagogical content knowledge via a PBL process.

Appendix B

Hamm Brothers Page One

In an Internet article, Josh Levin (2004c) has the following to say:

(Inserted here was the complete text of the editorial, *The 2004 Olympics: Why the U.S. men's gymnastics team is un-American*, which can be accessed online at http://www.slate.com/id/2105271/.)

The Hamm twins and their father responded that they are not in fact identical twins, but fraternal twins. What is the difference between identical and fraternal twins and does it make any difference to the argument of Mr. Levin?

Hamm Brothers Page Two

In a follow-up article the next day, Mr. Levin (2004a) continues with his discussion:

(Inserted here was the complete text of the editorial, *The 2004 Olympics: Are Paul and Morgan Hamm identical or fraternal twins?*" which can be accessed online at http://www.slate.com/id/2105357/.)

Questions:

1. What is meant by the terms *monozygotic* and *dizygotic*?
2. Why is a DNA test the only way to tell if they are identical or fraternal twins?
3. If they are identical, shouldn't they be the same in all respects? If they aren't, does this prove they are not identical?

Hamm Brothers Page Three

In the next day's column (Levin, 2004b), the following information was given:

(Inserted here was the complete text of the editorial, *The 2004 Olympics: Are Parul and Morgan Hamm identical twins? Their parents say no. Science says yes,"* which can be accessed online at http://www.slate.com/id/2105372/.)

Questions:

1. Does this article answer any of the questions that you asked or that you were asked before?
2. Do we now know if the Hamm brothers were identical or not?
3. Would it matter to their participation in the Olympics if they were identical? Would it be unfair to the other competitors if they were competing against identical twins, triplets or quadruplets?
4. If humans were cloned, they would be no more identical than identical twins. What if an exceptional athlete was cloned to produce ten or more copies? What would you think then?

Hamm Brothers Page Four

Big ideas.

1. Twins come from two eggs that were fertilized at about the same time.
2. Eggs may at times split into two independent cells either before or after fertilization.
3. Not all characteristics of twins are due to the genes they carry. Most characteristics result from the interplay of genes and the environment in which development and growth occurs.
4. Since even monozygotic twins do not develop and grow in identical environments, they are never identical in all respects; just more similar than fraternal twins or other siblings.

Michigan benchmarks.

Heredity (LH) III.3, Middle School

1. Describe how the characteristics of living things are passed on through generations.
2. Describe how heredity and environment may influence/determine characteristics of an organism.

Heredity (LH) III.3, Middle School

1. Explain how characteristics of living things are passed on from generation to generation.
5. Describe how genetic material is passed from parent to young during sexual and asexual reproduction (MDE, 2000).

Possible misconceptions.

1. Identical twins are identical in every respect.
2. All characteristics of individuals are determined by genes or DNA.
3. Cloning of multicellular organisms is unnatural.

Appendix C

Blood Doping Page One

Bicyclist Tyler Hamilton, one of Lance Armstrong's main rivals, was banned from competition last fall when he failed a blood test for *Blood Doping*, an illegal performance-enhancing procedure. He vehemently denied the charges, claiming that he was "100% innocent." In the weeks and months that followed, a debate raged in the media. Some experts claimed that the blood test was done sloppily; others that it wasn't a reliable test to use in the first place because several other factors could contribute to a false positive result; others that the most likely explanation was that Hamilton was, in fact, blood doping (B.B.C., 2004).

Shelia has students in her class who are following this debate in the media, taking emotionally charged positions on Tyler Hamilton's guilt or innocence. How can she use the debate to respond to her students and to engage them in the study of science?

Blood Doping Page Two

Blood doping is a procedure that can occur in one of several ways:

- An athlete can have between 1 and 4 units of the athlete's blood withdrawn several weeks before a key competition. Then the blood is centrifuged to remove the cells, and the plasma is immediately returned to the athlete's body. The cells can be refrigerated or frozen. They are transfused into the athlete usually 1 to 7 days before a high endurance event.
- The same procedure can be done with donated blood.
- An athlete can be injected with EPO. EPO is a genetically-engineered version of a natural hormone made by the kidney that stimulates bone marrow to make red blood cells.

Synthetic EPO is sold as a rescue medicine for treating anemia in end-stage kidney disease, when production of EPO declines.

Any of these blood doping methods are difficult to detect. There is no toxic substance for which to test the urine or blood, such as with steroid use. There are some methods currently being used for detection, but many are controversial (Beckham, n.d.; Sullivan, 1999).

Blood Doping Page Three

Full texts of two articles (Beckham, n.d.; Sullivan, 1999) were given to participants.

Blood Doping Page Four

Benchmarks from the *Michigan Curriculum Frameworks* (MDE, 2000)*:*

MCF LC III.1 m2: Explain why and how selected specialized cells are needed by plants and animals.

MCF LC III.1 h2: Compare and contrast ways in which selected cells are specialized to carry out particular life functions.

MCF LO III.2 m4: Explain how selected systems and processes work together in animals.

Appendix D

Length of Day Page One

One hundred fifteen miles south of the Arctic Circle, Fairbanks is home to people, animals, and plants that experience a wide range of temperature. People enjoy warm summer days yet have to bundle up in parkas to step outside of their snug, well-insulated homes in winter. Animals spend their summer days filled with activity, eating and busy with the work of raising their young. Fur coats become light in weight and dark in color. In winter, many animals migrate or spend their time in hibernation. Those that do brave the cold temperature often wear a coat of white. Plants grow quickly and abundantly in the warm summer, often completing their entire life cycle in a few short months. During the cold winter, only the conifers keep their green leaves.

What is driving this brutal change in temperature from summer to winter? (Participants were then given Table D1.) <Insert Table D1 about here.>

Length of Day Page Two

Lansing, Michigan is located 1660 miles south of the Arctic Circle. Lansing is also home to many people, animals, and plants. People living in Lansing experience a wide range of high temperatures throughout the year, but not of the magnitude of Fairbanks. Animals and plants also have adaptations to deal with the change in temperature, but they're not as extreme as those in Fairbanks.

Why isn't the change in temperature from summer to winter as extreme in Lansing as it is in Fairbanks? (Participants were then given Table D2.) <Insert Table D2 about here.

Length of Day Page Three

Resources.

Temperature data for any city is available at *The Weather Underground, Inc.* (2006) *website at* www.wunderground.com. Click on the monthly calendar link.

Sunrise and sunset data for any city can be found at the U.S. Naval Observatory website (2006), at http://aa.usno.navy.mil/data/docs/RS_OneDay.html.

Sunrise and sunset data for each week can be found at the *Mystery class in the journey north* website (Journey North, 2005) at www.learner.org/jnorth/mclass/spring2005. Click on "Updates" and follow the links to the sunrise/sunset data.

Globes, models of the sun and earth, and solar motion demonstration kits are available for you to use as resources.

The *Astronomy Society of the Pacific* website (2006) is another resource. It is found at www.astrosociety.org.

Length of Day Page Four

Big ideas.

- Seasons are defined by the amount of light in a day and whether it is increasing or decreasing with time.
- The tilt of the earth's axis causes the northern and southern hemispheres to have different amounts of light.
- The revolution of the earth around the sun causes the seasons to change.
- The tilt of the earth causes the sun to be high or low in the sky at noon.
- The angle of the sun in the sky causes the light hitting earth to be either concentrated or spread out.

- The total amount of light that hits a given location depends on the length of the day and the intensity of the light. The earth absorbs light energy and transforms it into heat energy, warming the air.

Benchmarks

- Compare and contrast the characteristics of the sun, moon, and earth. (EL #1)
- Describe the motion of the earth around the sun & the moon around the earth. (EL#2)
- Compare the earth to other planets and moons in terms of supporting life. (MS #1)
- Describe, compare, and explain the motions of solar system objects. (MS #2)
- Describe common observations of the day and night skies. (MS #3)
- Describe weather conditions and climates. (EL #1)
- Describe the composition and characteristics of the atmosphere. (MS #2)
- Describe seasonal changes in Michigan's weather. (EL #2)
- Explain patterns of changing weather and how they are measured. (MS #1)
- Explain the behavior of water in the atmosphere. (MS #3)
- Describe patterns of air movement and how they affect weather conditions. (HS #2)
- Explain and predict general weather patterns and storms. (HS #3)
- Explain the impact of human activities on the atmosphere. (HS #4

Misconceptions.

- The seasons are caused by variations in the earth's distance from the sun – summer is when the earth is closer to the sun and winter is when the earth is farther away.
- The tropics have longer days than the temperate zone in summer.
- The hottest day is the first day of summer.
- The hottest time each day is noon.

- The sun is directly over your head at noon each day.
- The sun follows the same path across the sky each day.
- The hottest temperatures occur in the tropics.
- The sun is always directly overhead (at a 90^o angle) at the equator.

Appendix E

Moon Phases Page One

The moon's shape seems to change from day to day, although we know that it is a solid sphere like the Earth (only smaller). What is happening to make the apparent shape of the moon change like this? Is there a pattern to the changes in the shapes we see? Does its position in the sky have anything to do with this?

Moon Phases Page Two

Participants were given pictures of the night sky in page two. There were 14 pictures, each taken at the same time each night for 14 consecutive days.

Moon Phases Pages Three and Four

This problem did not include a page three or a page four.

Table 1.
Criteria for Evaluating Written Content Dilemmas

Characteristic	**Descriptor**
Authentic and relevant	The problem is one that participants are reasonably likely to encounter and that they actually have a role in solving (Herreid, 1997; Kelson, 2004). Authentic tools of the discipline are available for participants' use in finding information towards the problem solution (Kelson, 2004).
Pedagogically effective	The problem naturally addresses the knowledge and skills for which it was designed, including content knowledge, problem solving, peer collaboration, and self-directed learning (Duch, 2001; Herreid, 1997; Kelson, 2004) The content that the problem addresses is sufficient to warrant the cost in time and effort. (Kelson, 2004).
Controversial	The problem is likely to provoke either multiple, viable hypotheses (Duch, 2001; Herreid, 1997; Kelson, 2004), or has competing, equally valid solutions (Harrington, 1995). The problem is ill-structured, with a "core of ambiguity" that can be interpreted in many useful ways (Hansen, 1997).
Appropriately complex	The problem provides no more or less information than one would naturally possess in the context of the problem, necessitates cooperation of all members (Duch, 2001; Kelson, 2004), requires rigorous academic work to solve, and requires participants to synthesize information from a variety of sources (Levin, 2001).
Clear about participants' roles	The roles of the participants flow naturally from the problem, and the product or performance required is clearly stated (Kelson, 2004). The problem specifies the role of the learner as a person who is a stakeholder in the solution of that particular problem (Torp and Sage, 2002), A decision-forcing moment exists, which engenders a desire to solve the dilemma and justify their decisions (Duch, 2001; Herreid, 1997).
Engaging	The problem stimulates curiosity, arouses interest, and motivates participants to probe for deeper understanding throughout the PBL process (Herreid, 1997; Kelson, 2004).

Table 2.
Participants' Ratings of the Highest and Lowest Rated Content Dilemmas in Two Break-out Groups in This Study

Names of Break-out Groups & Dilemmas	**Mean**	**N of Participants**
Astronomy (from Earth/Space Science mega-strand)		
Length of Day in AK & MI	1.88	8
	1.64	14
Moon Phases	2.00	8
Cells (from Life Science mega-strand)		
Hamm Brothers	0.60	5
Blood Doping	1.50	5

Table 3.
Ratings of the problems according to the criteria from the literature

	Cell Biology Problems		Astronomy Problems	
	Hamm Brothers	Blood Doping	Length of Day	Moon Phases
Authentic	0	2	1	1
Pedagogically effective	0	2	1	1
Controversial	0	2	0	0
Appropriately complex	0	1	1	1
Clear about participants' roles	0	2	0	0

Table D1
Data on Fairbanks, AK from the Length of Day dilemma

Month & Day	**Hours of Daylight**	**Sunrise**	**Sunset**	**Average high**	**Average low**	**Monthly precipitation (inches)**
June 21	21 hr 49 min	1:58 am	11:47 pm	72°F	52°F	1.74
July 21	19 hr 24 min	3:14 am	10:38 pm	72°F	53°F	2.14
August 21	15 hr 50 min	4:58 am	8:48 pm	62°F	46°F	2.27
September 21	12 hr 22 min	6:32 am	6:45 pm	51°F	35°F	1.38
October 21	9 hr 2 min	8:04 am	5:06 pm	28°F	12°F	0.94
November 21	5 hr 41 min	9:45 am	3:26 pm	14°F	-4°F	0.75
December 21	3 hr 43 min	10:58 am	2:41 pm	6°F	-12°F	0.68
January 21	5 hr 48 min	10:09 am	3:57 pm	1°F	-15°F	0.68
February 21	9 hr 9 min	8:31 am	5:40 pm	8°F	-18°F	0.5
March 21	12 hr 24 min	6:47 am	7:11 pm	29°F	2°F	0.48
April 21	15 hr 52 min	4:53 am	8:45 pm	50°F	27°F	0.32
May 21	19 hr 24 min	3:07 am	10:31 pm	59°F	39°F	0.74

Table D2
Data on Lansing, MI from the Length of Day dilemma

Month & Day	**Hours of Daylight**	**Sunrise**	**Sunset**	**Average high**	**Average low**	**Monthly precipitation (inches)**
June 21	15 hr 20 min	6:00 am	9:20 pm	79°F	56°F	3.7
July 21	14 hr 51 min	6:19 am	9:10 pm	82°F	59°F	2.5
August 21	13 hr 39 min	6:51 am	8:30 pm	79°F	55°F	3.2
September 21	12 hr 12 min	7:25 am	7:37 pm	66°F	47°F	3.6
October 21	10 hr 47 min	7:59 am	6:46 pm	56°F	37°F	2.1
November 21	9 hr 32 min	7:38 am	5:10 pm	42°F	27°F	2.6
December 21	9 hr 3 min	8:05 am	5:08 pm	31°F	17°F	2.3
January 21	9 hr 36 min	8:02 am	5:38 pm	30°F	14°F	1.5
February 21	10 hr 52 min	7:26 am	6:18 pm	36°F	19°F	1.4
March 21	12 hr 13 min	6:39 am	6:52 pm	43°F	25°F	2.3
April 21	13 hr 40 min	6:47 am	8:27 pm	60°F	40°F	2.8
May 21	14 hr 50 min	6:10 am	9:00 pm	72°F	46°F	2.6