Physics Problem Solving

Jennifer L. Docktor

University of Minnesota

Abstract

Problem solving is viewed as a fundamental part of learning physics, and research to improve the teaching and learning of physics problem solving is a primary subfield of Physics Education Research (PER). A difficult question in this field has been how to measure problem solving performance. Answering this question requires a definition of what problem solving is, and what successful problem solving looks like. Existing strategies to assess physics problem solving that focus on grading the "correctness" of a solution give an inadequate description of students' skills, and other strategies based on criteria or ranking scales are in need of further testing and development. In this paper, I present a review of literature on problem solving and efforts to measure problem solving performance, including those at the University of Minnesota.

Table of Contents

Physics Problem Solving

1: Introduction

Problem solving is viewed as a fundamental part of learning physics (Heller, Keith, & Anderson, 1992; McDermott, 1981; Reif, 1981; Reif, Larkin, & Brackett, 1976). However, many physics instructors find that their students do not solve problems at the desired level of proficiency (Redish, Scherr, & Tuminaro, 2006; Reif, 1995; Van Heuvelen, 1991). To help improve the teaching and learning of physics problem solving, studies commenced in the 1970's and early 1980's to understand these difficult tasks (McDermott & Redish, 1999). Since that time, systematic research into physics problem solving has become recognized as a primary subfield of Physics Education Research, or PER (Hsu, Brewe, Foster, & Harper, 2004).

Physics Education Research has grown substantially in the past few decades to include several topics of study, such as: students' alternative conceptions about physics topics, problem-solving performance, laboratory instruction, lecture demonstrations, mathematical proficiency and its application to physics, and the attitudes and beliefs of students about learning physics (McDermott & Redish, 1999). Most of the empirical studies into these areas of research have focused on students' conceptions, and on problem-solving performance. The latter has been an explicit focus of the PER group at the University of Minnesota, and subsequently will be the center of attention for this review.

The capacity to solve complex problems is considered an essential skill for citizens of today's changing technological society (Martinez, 1998). In fact, some institutions list problem solving as a desired learning outcome for all undergraduate students, regardless of their chosen area of study. In October 2006 Vice Provost Arlene Carney reported (Carney, 2006) that the Provost's Council for Enhancing Student Learning proposed seven undergraduate learning outcomes for all University of Minnesota students, the first of which is, "At the time of receiving a bachelor's degree, students will demonstrate…the ability to identify, define, and solve problems." The newsletter goes on to expand its description of this goal in a chart with the elaboration: University of Minnesota graduates recognize the complexity and ambiguity inherent in many problems, can evaluate and synthesize knowledge and frame logical arguments based on this knowledge, and understand and use the scientific method and other modes of problem solving (Carney, 2006).

The context of physics presents a good opportunity for students to engage in problem solving. As well as learning major concepts and principles of physics, problem solving skills are considered a primary goal of physics instruction, both in high school and college physics courses (Hsu et al., 2004; Redish et al., 2006; Reif et al., 1975). A survey of science and engineering faculty at the University of Minnesota suggests that a primary reason departments require their undergraduate students to take introductory physics is to learn quantitative and qualitative problem solving skills (Heller, Heller, & Kuo, 2004).

Since it has been established that problem solving skill is a goal valued by society, institutions of higher education, and physics courses, how do we determine progress toward that goal? For example in the recent NRC report "Rising Above the Gathering Storm" (Augustine, 2006) the statement is made:

> The No Child Left Behind legislation requires testing of students' knowledge of science beginning in 2006-2007, and the science portion of the NAEP is being redesigned. Development of such assessments raises profound methodologic issues such as how to assess inquiry and problem-solving skills using traditional large scale testing formats (p. 264)

Answering this question requires a definition of what problem solving is, and what successful problem solving looks like. These topics will be addressed in the following section. An evaluation of problem solving skill also requires a means by which to "measure" this complex cognitive process, which is a non-trivial task. Efforts to assess problem solving will be addressed in section three.

2: Research Literature on Problem Solving

The research literature on problem solving extends into several disciplines, including cognitive science, psychology, education, mathematics, and physics (Hsu et al., 2004). References about research on problem solving can be sorted into several categories, including: definitions of "problem" and "problem solving", cognitive aspects of problem solving (such as use of representations, strategies, knowledge structures, and cognitive load and working memory), expert-novice characteristics of procedures and knowledge structures, problem solving in mathematics and its transfer to other contexts, alternative problem types, curricular interventions, and computers and problem solving. For the purposes of this literature review, the first three of these categories will be explored in greater detail.

2.1: Definitions of "problem" and "problem solving"

Attempts to describe what are meant by a "problem" and by "problem solving" in the research literature are surprisingly consistent. For example, Newell and Simon (1972) write that, "A person is confronted with a *problem* when he wants something and does not know immediately what series of actions he can perform to get it" (p. 72). Similarly, Martinez (1998) states, "problem solving is the process of moving toward a goal when the path to that goal is uncertain" (p. 605). The mathematician Pólya (1962) defines problem solving as a search "for some action appropriate to attain a clearly conceived, but not immediately attainable, aim," and adds, "where there is no difficulty, there is no problem" (p. 117).

One commonality within these definitions is their subjectivity. What is considered a "problem" for one person may not be a problem for another person – the definition depends on the perceived difficulty of the task (Hsu et al., 2004). Problem solving also depends on a person's prior experience; some skills can be learned well enough that they become automatic and require minimal effort (Ormrod, 2004).

The basic components of a problem are considered to be information provided or an initial state (*givens*), a desired end state (*goal*), and means to get from the initial state to the end state, (*operations)* (Ormrod, 2004). Problems can differ vastly, however, in their structure. On one extreme of the continuum are problems that are straightforward, that clearly state the givens and desired goal, and for which all information needed to solve the problem "correctly" is presented. These are referred to as well-defined problems (Ormrod, 2004; Pretz, Naples, & Sternberg, 2003). On the other extreme are problems for which the desired goal could be uncertain, some necessary information is absent, and for which there might be several possible solutions. These are termed ill-defined problems.

The degree to which a problem is considered well-defined or ill-defined depends on an individual's expertise, and therefore solvers will differ in their problem-solving approaches and strategies. For expert problem solvers, it is possible that they will be presented with a problem and know immediately the steps to use in solving it. For this situation, some researchers make the distinction that it is not a "problem" for them at all, but rather can be classified as an "exercise".

Cognitive psychologists describe seven steps that comprise a problem solving cycle (Pretz et al., 2003). They assert that a problem solver must first recognize the existence of a problem and identify it, define and represent the problem, develop a strategy or plan to reach a solution, organize his or her knowledge, allocate mental and physical resources to the problem, monitor progress toward the goal, and evaluate the solution.

All steps may not be traversed for every kind of problem. For example, when a problem is presented to the solver directly, the first step of identifying the existence of a problem might not be required. Even when a problem is presented, however, the solver must still interpret the problem statement and redefine it in terms of their personal understanding. In contrast, problems that are discovered or created in the course of another task almost always require extensive definition and representation.

How you identify, define, and represent a problem is mediated by a variety of internal and external factors. Internally, existing knowledge and expectations frame the interpretation of a problem and mediate the capability to represent the problem efficiently (Pretz et al., 2003). Other internal factors such as individual differences in abilities and dispositions can also affect the problem solving process; for example, persons with high spatial abilities are more likely to represent problems with the use of images rather than linguistically. The external factor of social context (peers, culture, and language) can affect the types of problems that are recognized by a group, and the terms used to describe a problem.

2.2: Theoretical frameworks from psychology

Early psychological theories to describe problem solving were based on the ideas of trial-and-error behavior and stimulus-response associations (Ormrod, 2004). From Edward

Thorndike’s classic observations of animals in puzzle boxes from the early 1900’s and additional observations of young children, psychologists concluded that trial-and-error behavior is a generally ineffective and time consuming approach to problem solving and is only workable if there is a limited number of possibilities to try. Based on these observations psychologists such as Thorndike, Pavlov, and Skinner developed the idea of conditioning based on stimulus-response connections, which focused on observable stimuli in the environment and an organism’s responses to those stimuli. This perspective of learning became known as behaviorism, because research was based on observable behavior that could be measured.

Research experiments by neobehaviorist Hull expanded the stimulus-response approach to consider factors such as motivation and the influence on the strength of a stimulus-response association on behavior (Ormrod, 2004). In his theory, an organism learns different responses for the same stimulus, and organizes them in what is called a “response hierarchy” according to habit strength. When an organism is presented with a problem (stimulus) the first response produced will have the strongest association with the problem, perhaps based on the perceived effectiveness of that response for solving similar problems in the past. If that response doesn’t work, other responses will be produced in order of their strength until either the problem is solved or the list of known responses is exhausted.

In early theories of learning and problem solving, psychologists focused their research only on processes that can be observed and measured, and therefore they largely excluded internal processes from studies (Ormrod, 2004). Although behaviorist and neobehaviorist approaches can be used to explain some aspects of human behavior such as problem solving, they have largely been abandoned in favor of the cognitive perspective. This shift in the field began with what is called Gestalt psychology, and then further developed into information processing theory, which remains a prevalent theory today.

A focus on mental processes involved in learning and problem solving emerged from Köhler's observations of chimpanzees (Ormrod, 2004). Rather than attribute their behavior to trial-and-error, Köhler thought that the chimps appeared to carefully examine all parts of a problem and deliberately take action in their attempts to solve the problem. He concluded that problem solving is a process of combining and recombining components of a problem mentally (what he called "restructuring") until a point of insight into the problem solution is reached. Other early cognitive approaches attempted to identify the steps or mental stages in the problem solving process, and some of these theories will described in section 2.3.

Contemporary theories of learning and problem solving are based on the information processing theory of memory (Ormrod, 2004; Redish, 2003). In brief, this theory asserts that information in the brain is stored in two primary components of memory: short term or "working" memory and long-term memory. Short-term memory is limited in size and the length of time it can hold information, and is thought to contain distinct verbal and visual parts (Redish, 2003). In contrast, long-term memory can hold vast quantities of facts and data for long periods of time, but to be accessed this information must be activated by being brought into working memory.

In problem solving, it is thought that working memory is utilized to process information about the problem and maintains its accessibility during the problem solving process. Since working memory has limited storage capacity, it is possible that information in a problem can exceed this limit and interfere with attempts to seek a solution. For this reason, information about a problem is often stored externally (written down) or processed externally (such as with a calculator or computer) in order to free up space in working memory that can be devoted to the task. Also, some skills involved in problem solving can be practiced until they become automatic, which also minimizes the use of working memory capacity.

In addition to utilizing short-term memory, problem solving also requires accessing relevant information - the solver's knowledge base about the problem - from storage in long-term memory (Ormrod, 2004). Critical factors in this retrieval of information include *what* has been stored and *how* it has been stored, and the cues or patterns present in the problem that help the individual perceive what information to access from memory. In order to be retrieved, knowledge from a content domain must be present in memory to begin with; it must also be organized in a way that facilitates its retrieval in an appropriate context.

With experience in a content domain, it is believed that problem solvers develop cognitive structures called "problem schemata" that allow them to recognize a problem as belonging to a particular category (Sweller, 1988). This mental classification of problem types can trigger particular actions for solving the problem, based on the perceived similarity of the presented problem to the same category of others stored in memory. There is evidence that expert and novice problem solvers differ in their problem schemata, which will be described in more detail in section 2.4.

The cognitive approach to problem solving also includes the concept of metacognition, which refers to an individual's awareness of his or her own thinking process (Martinez, 1998; Ormrod, 2004). Metacognitively engaged problem solvers have developed skills at planning their problem solving approach, monitoring their progress toward the goal while following their plan, and evaluating the effectiveness of their chosen strategies. Since metacognitive solvers are careful to evaluate their assumptions and are less apt to persevere in unproductive strategies, they are more likely to solve complex problems successfully.

2.3: General and specific strategies for solving problems

Descriptions of operations or methods for solving problems often draw a distinction between *algorithms* and *heuristics* (Martinez, 1998; Ormrod, 2004; Pretz et al., 2003). The term algorithm is usually applied to step-by-step procedures that will guarantee a correct solution every time, if applied correctly. An example of an algorithm is the mathematical procedure for carrying out long division (Ormrod, 2004). The term heuristic is used to refer to general strategies or "rules of thumb" for solving problems (Martinez, 1998; Ormrod, 2004). Examples of heuristics include combining algorithms, hill climbing, successive refinements, means-ends analysis, working backward, and using visual imagery or external representations.

Basic-level mathematics and physics problems are often solved using the heuristic of combining algorithms, by using several algebraic procedures in succession (Ormrod, 2004). In hill climbing, problem solvers perform actions that bring them closer to the goal, and in successive refinements, they solve a simpler related or idealized problem that will help them understand the presented problem. Successive refinements is a process common in writing; initially a rough outline or draft is produced, and then it is gradually revised to a finished product. Means-ends analysis is perhaps one of the most well-known heuristics, by which a solver breaks the problem up into subgoals, and then works successively on each of them (Martinez, 1998; Ormrod, 2004). Another common heuristic is working backward, in which the solver starts at the desired goal and considers the reasonable steps just prior to that goal, and continues in this process until reaching the initial state. The use of external representations or visual imagery is a means to represent the complexity of a problem and free up some space in working memory; it also creates a public document that can be examined by others.

As cited in Ormrod (2004), one general problem solving strategy includes defining the problem and gathering information relevant to its solution (*preparation*), thinking about the problem at a subconscious level while engaging in other activities (*incubation*), having a sudden insight into the solution of the problem (*inspiration)*, and checking to be certain that the solution is correct (*verification*). A difficulty with the usefulness of this strategy, however, is that it is not clear how to facilitate the occasion of inspiration, and extended periods of incubation are not always feasible.

The mathematician Pólya (1957) is also often cited for his 4-step problem solving strategy. His first step is *Understanding the Problem*, by identifying the unknown, the data, and the condition, and then drawing a figure and introducing suitable notation. The second step is *Devising a Plan*, in which the solver seeks a connection between the data and the unknown. If an immediate connection is not found, the solver considers related problems or problems that have already been solved, and uses this information to devise a plan to reach the unknown. In the third step, *Carrying Out the Plan*, the steps outlined in part two are carried out, and each step is checked for correctness. In the final step *Looking Back*, the problem solution is examined, and arguments are checked.

Strategies that are specific to physics have been developed by Reif (1995) in his textbook *Understanding Basic Mechanics*, and by Heller and Heller at the University of Minnesota (Heller & Heller, 2000; Redish, 2003). Reif's steps include *Analyze the Problem*, in which a basic description of the situation and goals is generated, and a refined physics description according to time sequences and intervals is developed (Reif, 1995). The second step is *Construction of a*

Solution, in which basic useful relations are identified and implemented until unwanted quantities are eliminated. The final step is called *Checks*, and asks the solver if the goal has been attained, the answer is in terms of known quantities, and there is consistency within the solution in terms of units, signs, and sensibility of values.

The steps of the University of Minnesota problem-solving strategy include *Focus the Problem*, which involves determining the question and sketching a picture, and selecting a qualitative approach (Heller & Heller, 2000; Redish, 2003). The next step, *Describe the Physics*, includes drawing a diagram, defining symbols, and stating quantitative relationships. The step *Plan a Solution* entails choosing a relationship that includes the target quantity, undergoing a cycle of choosing additional relationships to eliminate unknowns, and substituting to solve for the target. The step *Execute the Plan* involves simplifying an expression, and putting in numerical values for quantities if requested. The final step is *Evaluate the Answer*, which means evaluating the solution for reasonableness, completeness, and to check that it is properly stated.

This process can be thought of as a series of translations, in which each step converts the previous step into a slightly different "language" (Heller, et al., 1992). For example, in *Focus the Problem* the words of the problem statement are translated into a visual representation in the form of a sketch. In the *Describe the Physics* step, the sketch is translated into a physical representation of the problem that includes a diagram and symbolic notation. *Plan a Solution* is a translation from the physics description into mathematical form using equations and constraints, which are further translated into mathematical actions to obtain an arithmetic solution in *Execute the Plan*.

2.4: Expert-novice characteristics

The research literature on problem solving has shown differences between experienced (or "expert") problem solvers and those who are inexperienced (or "novices") both in their procedures for solving problems and their organization of knowledge in memory (Chi, Feltovich, & Glaser, 1980; Larkin & Reif, 1979). In physics problem solving, novice students tend to spend little time representing the problem and quickly jump into quantitative expressions (Larkin, 1979; Pretz et al., 2003). Instructors have found that novice students implement problem solving techniques that include haphazard formula-seeking and solution pattern matching (Reif et al., 1976; Mazur, 1997; Van Heuvelen, 1991).

In contrast, experts solve problems by interjecting an additional step of a qualitative analysis or a low-detail overview of the problem before writing down equations (Larkin, 1979; Pretz et al., 2003). This qualitative analysis used by experts, such as a verbal description or a picture, serves as a decision guide for planning and evaluating the solution (Larkin & Reif, 1979). Although this step takes additional time to complete, it facilitates the efficient completion of further solution steps and in most cases the expert is able to successfully complete the problem in less time than a novice (Pretz et al., 2003). This description also often explores the constraints inherent in the problem, such as expectations for extreme values of parameters in the problem, which can aid in a final check or evaluation of the problem solution (Reif & Heller, 1982).

In addition to differences in procedures, experts and novices differ in their organization of knowledge about physics concepts. Larkin (1979, 1980) suggested that experts store physics principles in memory as "chunks" of information that are connected and can be usefully applied together, whereas novices must inefficiently access each principle or equation individually from memory. As a result of this "chunking" of information, the cognitive load on an expert's short-term memory is lower and they can devote more memory to the process of solving the problem (Pretz et al., 2003; Sweller, 1988). For a novice, accessing information in pieces places a higher cognitive load on short-term memory and can interfere with the problem solving process.

Chi et al. (1980) found that experts categorize physics problems based on underlying structure or physics principles involved, whereas novices look at the surface features of the problem such as the objects mentioned in the problem description. They further hypothesized that these categorizations indicate that the problem schemata of experts and novices contain different knowledge which influence their problem representations and the approaches used by those experts and novices.

3: Problem Solving Rubric

3.1: Problem solving assessment strategies

How is problem solving performance measured? In most introductory physics courses, students' problem solutions on homework or exams are given a score based on the correctness of the algebraic or numerical solution (Heller et al., 1992). A standard grading practice in physics involves giving students partial credit for particular characteristics of their written solution, as compared to the ideal solution developed by the instructor. Simply comparing average scores based on this grading scheme, however, does not give an adequate description of the student's problem solving performance. At best it only gives an indication of whether one solution is "better" than another in terms of the prescribed grading scheme. A different kind of instrument is required to determine the nature of a student's approach to the problem and assess a solution in terms of characteristics of "expertise" in solving problems.

Research into problem solving has used several different means to measure problem solving performance. One method used by Larkin and Reif (1979) involves measuring the time it takes a problem solver to write down each quantitative expression in their solution, and recording the total time to reach a solution. Some researchers have also investigated problem solving using think-aloud protocols or interviews, in which students engage in conversation to explain their thinking processes as they attempt a physics problem. A difficulty with these methods is the time involved to prepare and conduct them, the vast amount of data generated from interview transcriptions, and the complicated nature of the data analysis. In order to compare problem solving performance for many students, it is desirable to have a quantitative measure that can be determined relatively quickly. Researchers who have attempted to assess problem solutions on the basis of criteria or characteristics include Reif and Heller (1982), Heller, Keith, and Anderson (1992) and the theses of Blue (1997) and Foster (2000).

Research by Reif and Heller (1982) assessed physics problem solutions on the basis of five criteria. The first criterion of *Clear Interpretation* requires that the parameters of the problem are clearly defined and specifications are provided (such as direction and magnitude of vectors, direction of motion, and reference frames). The criterion of *Completeness* requires that all questions are answered completely and expressed only in terms of known quantities. *Internal Consistency* means that the solution is void of logical errors, and *External Consistency* means that the answers agree with expected relationships between parameters. The final criterion of *Optimality* requires that the solution is simple and easily interpretable.

An investigation described by Heller, Keith, and Anderson (1992) uses a rating scale of problem solving performance based on six characteristics of expert-like problem solutions. These characteristics include evidence of conceptual understanding, usefulness of the problem description, consistency of the specific equations with the physics description written, reasonableness of the plan, logical progression of the mathematical solution from physics principles to problem-specific expressions, and the use of appropriate mathematics. The characteristics in this scheme are weighted equally and normalized to obtain a score out of 100 points.

Similar coding rubrics were used to measure problem solving in the unpublished doctoral dissertations of Jennifer Blue (1997) and Tom Foster (2000) at the University of Minnesota. Although considerable effort was put forth to develop and test the rating criteria used in these theses, a problem solving rubric is still in the developmental phase. Further testing and revisions are necessary to meet desired specifications from researchers in Physics Education and instructors of physics courses (faculty). Some of the important design considerations for a problem solving rubric are discussed in the next section, and progress on the development of such a rubric at the University of Minnesota will be discussed in section 3.3.

3.2: Constraints and considerations

An evaluative tool to assess problem solving performance must take into consideration its intended use in both research and instruction. In the case of physics education research, a problem solving rubric is useful to rank the problem solving performance of students as repeatedly measured over time, such as throughout a single semester. Having a common ranking scale can also facilitate comparison between students' problem solving for different courses or even across institutions. As a result, the criteria in the rubric must be described in enough detail that its use to evaluate a problem solution can be easily learned. The scores obtained by different people must also be reliable (inter-rater reliability).

In the case of instruction, a problem solving rubric can be used as an alternative to standard grading schemes. Since instructors typically cover a wide variety of topics in their courses, a problem solving rubric must be general enough that it applies to a range of problem topics. If the intent is to replace a standard grading scale for homework or exams, the criteria must be consistent with instructors' values with respect to "good" or "bad" problem solutions.

3.3: Description of a problem solving rubric for physics

As stated previously, work on developing a problem solving rubric at the University of Minnesota (U of MN) began with the investigations by Heller, Keith, and Anderson (1992) and the doctoral dissertations of Jennifer Blue (1997) and Tom Foster (2000). Since that time, the coding scheme has been modified over several iterations from 2004 to 2006 by graduate students and a post doc working with the Physics Education Research group during that time period. The revision process involved coding example student solutions to introductory course final exam problems using the rubric, and discussing scores given by each rater. The most recent version of the rubric is in Appendix 1.

The criteria of the U of MN rubric are based on characteristics of expert and novice problem solutions described in the problem solving research literature. Earlier schemes that included six or more dimensions have been condensed into four dimensions: Physics Approach, Translation of the Physics Approach, Appropriate Mathematics, and Logical Progression. Each of these categories is described in more detail on the first page of the rubric in Appendix 1.

Each category of the rubric contains a list of six or seven numbered levels, where a lower score represents more novice-like solution characteristics and a higher score describes criteria for expert-like solutions. For example, novice-like characteristics could include fundamental errors in physics principles, logic, or mathematics. Expert-like characteristics could include appropriate use of physics principles, consistent progress toward a solution, and correct use of mathematics. A score of zero usually means that the solution cannot be coded; either the solution does not exhibit any criteria of that dimension or is too incomplete to be interpreted by the coder.

The U of MN problem solving rubric is still in the developmental phase, so further testing with example student solutions by multiple coders should continue. In particular, it would be desirable to have the same number of levels for each dimension, so this is a revision goal. The language of the criteria must also be easily understood by different people, so pilot testing with other coders not in the group could contribute further revisions. Once there has been reasonable agreement on the rubric among physics education researchers, it should be tested for inter-rater and intra-rater reliability using appropriate statistical measures. The instrument should also be examined for consistency with physics instructors' values for problem solving, so that it is more likely to be viewed as a useful tool by instructors and considered for implementation in their own scoring of students' homework and exams.

References

Augustine, N. (2006). *Rising Above The Gathering Storm: Energizing and Employing America for a Brighter Economic Future.* Washington, DC: National Academies Press.

Blue, J. M. (1997). *Sex differences in physics learning and evaluations in an introductory course.* Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Carney, A. (2006, October). What do we want our students to learn? *Transform*, *1*, 1-6.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1980). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, *5*, 121-152.

Foster, T. (2000). *The development of students' problem-solving skills from instruction emphasizing qualitative problem-solving*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 176-206). Cambridge, UK: Cambridge University Press.

Heller, J. I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction*, *1*(2), 177-216.

Heller, K., & Heller, P. (2000). *The competent problem solver for introductory physics*. Boston: McGraw-Hill.

Heller, P., Heller, K., & Kuo, V. (2004, January). *Procedure for setting goals for an introductory physics course*. Contributed talk presented at the meeting of the American Association of Physics Teachers, Miami Beach, FL.

Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, *60*(7), 627-636.

Hsu, L., Brewe, E., Foster, T. M., & Harper, K. A. (2004). Resource letter RPS-1: Research in problem solving. *American Journal of Physics*, *72*(9), 1147-1156.

Larkin, J. H. (1979). Processing information for effective problem solving. *Engineering Education*, *70*(3), 285-288.

Larkin, J. H. (1980). Teaching problem solving in physics: the psychological laboratory and the practical classroom. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: issues in teaching and research* (pp. 111-125). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, J. H., & Reif, F. (1979). Understanding and teaching problem solving in physics. *European Journal of Science Education*, *1*(2), 191-203.

Martinez, M. E. (1998). What is Problem Solving? *Phi Delta Kappan*, *79*, 605-609.

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.

McDermott, L. C. (1991). Millikan Lecture 1990: What we teach and what is learned – Closing the gap. *American Journal of Physics*, *59*(4), 301-315.

McDermott, L. C., & Redish, E. F. (1999). Resource letter: PER-1: Physics education research. *American Journal of Physics*, *67*(9), 755-767.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Ormrod, J. E. (2004). *Human learning* (4th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Pólya, G. (1957). *How to solve it* (2nd ed.). Princeton, NJ: Princeton University Press.

Pólya, G. (1962). *Mathematical discovery*. New York: John Wiley & Sons, Inc.

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 3-30). Cambridge, UK: Cambridge University Press.

Redish, E. F. (2003). *Teaching physics with the physics suite*. Hoboken, NJ: Johns Wiley & Sons, Inc.

Redish, E. F., Scherr, R. E., & Tuminaro, J. (2006). Reverse engineering the solution of a "simple" physics problem: Why learning physics is harder than it looks. *The Physics Teacher*, *44*(5), 293-300.

Reif, F. (1981). Teaching problem solving – A scientific approach. *The Physics Teacher*, *19*(5), 310-316.

Reif, F. (1995). Millikan Lecture 1994: Understanding and teaching important scientific thought processes. *American Journal of Physics*, *63*(1), 17-32.

Reif, F. (1995). *Understanding basic mechanics*. New York: John Wiley & Sons, Inc.

Reif, F., & Heller, J. I. (1982). Knowledge structures and problem solving in physics. *Educational Psychologist*, *17*(2), 102-127.

Reif, F., Larkin, J.H., & Brackett, G.C. (1976). Teaching general learning and problem-solving skills. *American Journal of Physics*, *44*(3), 212-217.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, *12*, 257-285.

Van Heuvelen, A. (1991). Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics*, *59*(10), 891-897.

Appendix 1: ***Problem Solving Rubric v. 4.4 July 13, 2006***

The following rubric is a coding guide to assess problem-solving performance. Ratings are based on characteristics of expert-like problem solutions in four distinct categories: Physics Approach (PA), Translation of Physics Approach (T), Appropriate Mathematics (AM), and Logical Progression (LP). This scoring design was modified from earlier schemes used by physics education researchers at the University of Minnesota.

Each category is subdivided into 5-6 levels designated by a numerical "score" in the left column and a description of the criteria met to attain that score in the right column. A score of zero usually means the solution is blank or too incomplete to be interpreted by the coder. A lower numbered score indicates that errors in the solution are more novice-like in nature (such as fundamental errors in physics principles, logic, or mathematics). Higher numbers describe minimal errors and expert-like thinking for that category. When using the rubric to code problem solutions, it is recommended to advance through the code descriptions for a category in ascending order and rate the solution on the first (lowest) score that describes the solution, even if it also matches statements of higher code numbers.

Physics Approach assesses an initial conceptual understanding of the problem statement. Typically this is indicated through an explicit description of the physics principles the student believes are necessary to solve the problem and a visual representation of the problem situation, such as a labeled sketch or physics diagram. An expert problem solver has an accurate interpretation of the problem statement and the physics principles needed to solve the problem and can describe them schematically, whereas a novice often has difficulty interpreting the problem statement, accessing the useful physics principles, and translating their ideas to situation representations.

Translation of Physics Approach assesses a student's success in converting stated physics principles and situation representations into symbolic form as specific physics equations. An expert problem solver has an extensive declarative and procedural knowledge base that facilitates efficient access to useful information for symbolic translation of their Physics Approach. A novice's knowledge structure is sparser and chunked differently; therefore we expect to find more serious mismatches between their Physics Approach and specific equations.

Appropriate Mathematics assesses the performance of mathematical operations on specific physics equations to isolate the desired quantity(s) and obtain a reasonable numerical answer without fundamental mathematical errors. An expert problem solver makes few math or calculation errors and substitutes numerical values for quantities at the last step, whereas a novice might violate rules of algebra, calculus, or geometry and make use of unjustified relationships in their attempt to reach a solution. If mathematical operations stop before a solution is reached, more expert-like students will do so because continuing will violate the rules of mathematics.

Logical Progression assesses the overall cohesiveness of a problem solution. An expert problem solver makes consistent progress from the problem statement to a physics description and specific equations that terminate in a reasonable answer. A novice might fail to progress to an answer, violate logic to reach an answer, or include significant unnecessary steps.

Physics Approach *(from problem statement to physics description)*

0	Nothing written can be interpreted as a physics approach.
1	All physics used is incorrect or inappropriate. Correct solution is not possible.
2	Use of a few appropriate physics principles is evident, but most physics is missing, incorrect, or inappropriate.
3	Most of the physics principles used are appropriate, but one or more principles are missing, incorrect, or inappropriate.
4	(Apparent) use of physics principles could facilitate successful solution; missing explicit statement of physics principles.
5	Explicit statement of physics principles could facilitate successful problem solution.

Translation of Physics Approach *(from Physics Approach to specific equations)*

0	Nothing written can be interpreted as a translation of the specified physics approach.
1	Fundamental error(s) in physics, such as treating vectors as scalars or not distinguishing between energy and momentum.
2	Missing an explicit symbol for a physics quantity or an explicit relationship among quantities essential to the specified physics approach.
3	Misunderstanding the meaning of a symbol for a physics quantity or the relationship among quantities.
4	Limited translation errors (i.e. sign error, wrong value for a quantity, or incorrect extraction of a vector component).
5	Appropriate translation with a correct but not explicitly defined coordinate system (such as x-y coordinate axes, current direction in a circuit, or clockwise/counterclockwise rotation).
6	Appropriate translation with explicitly defined coordinate system.

Appropriate Mathematics *(from specific equations to numerical answer)*

0	Nothing written can be interpreted as mathematics.
1	Mathematics made significantly easier (than a correct solution) by inappropriate translation from problem statement to specific equations.
2	Solution violates rules of algebra, calculus, or geometry. "Math-magic" or other unjustified relationship produces an answer (with "reasonable" units, sign, or magnitude).
3	Careless math or calculation error or unreasonable answer or answer with unknown quantities
4	Mathematics leads from specific equations to a reasonable answer, but features early substitution of non-zero numerical values for quantities.
5	Appropriate mathematics with possible minor errors (i.e. sign error or calculator error) lead from specific equations to a reasonable answer with numerical values substituted in the last step.

Logical Progression *(entire problem solution)*

0	Nothing written can be interpreted as logical progression.
1	Haphazard solution with obvious logical breaks.
2	Part of the solution contradicts stated principles, the constraints shown in the explicit statement of the problem situation, or the assumptions made in another part of solution.
3	Solution is logical but does not achieve the target quantity or haphazard but converges to the target.
4	Solution is organized but contains some logical breaks.
5	Progress from problem statement to an answer includes extraneous steps.
6	Consistent progress from problem statement to answer.