

A Structural Equation Model of Expertise in College Physics

Gita Taasoobshirazi
University of Nevada, Las Vegas

Martha Carr
University of Georgia

A model of expertise in physics was tested on a sample of 374 college students in 2 different level physics courses. Structural equation modeling was used to test hypothesized relationships among variables linked to expert performance in physics including strategy use, pictorial representation, categorization skills, and motivation, and these variables were examined for their influence on physics achievement. Gender was included in the model to examine how it influenced achievement indirectly through its influence on the other variables in the model. Two levels of expertise were examined by testing the model on trigonometry-based physics students and on more advanced, calculus-based physics students. Results were similar across both levels of expertise: For both courses, student motivation had a significant influence on students' strategy use and categorization skills. Categorization skills, in turn, influenced student achievement directly, and indirectly, through strategy use. Strategy use had a significant influence on achievement. Pictorial representation played little role in the model. Gender contributed primarily through motivation, but for the more advanced level course it also directly predicted strategy use.

Keywords: college physics, expertise, gender, problem solving, structural equation modeling

A Framework of Expertise in Physics

A number of critical variables discriminate experts from novices. Within their domain of expertise, experts use more goal directed strategies for solving problems; have greater knowledge, more organized knowledge, and greater motivation; tend to receive more social support; and are better at monitoring their performance (e.g., Alexander, 2003; Ericsson, 2006; Hatano & Oura, 2003; Zimmerman, 2006). Within the domain of physics, the bulk of the research comparing experts and novices focuses on three main differences. This research indicates that the way physics problems are solved, including the actual problem-solving strategies, the way problems are categorized, which suggests depth, breadth, and organization of conceptual knowledge, and whether problems are pictorially represented distinguish novices from experts (e.g., Anzai, 1991; Chi, 2006; Taasoobshirazi & Carr, 2008).

Although research specifies the variables that characterize expertise in physics, there is a lack of published research examining the relationship among these variables and a lack of research indicating which variables are most important for expert performance. As a result, physics instructors do not know what should be focused on during instruction and what is of less importance. For instance, although it might be expected that the conceptual knowledge used to expertly categorize and solve problems is the most important variable for expertise in physics, in practice, much more focus is placed on the importance of pictorial representations (Cutnell & Johnson, 2007). Given that high school-level and

college-level physics instruction often leaves students confused about basic concepts of mechanics, optics, thermodynamics, electricity, and magnetism and tends to lead to the use of poor problem-solving strategies (e.g., Henderson, 2005; McDermott, 2001), researchers need to better understand what teachers should focus on during instruction.

Females from kindergarten to 12th grade (K-12) and in college appear to have a hard time obtaining expert levels of performance in physics. The largest gender differences in achievement and participation exist in physics (Institute of Education Sciences, 2007; Murphy & Whitelegg, 2006; National Science Foundation [NSF], 2004). While the research on gender differences in science achievement has focused primarily on differences in motivation as a possible explanation (e.g., Halpern, 2000; Haussler & Hoffman, 2002; Morgan, Isaac, & Sansone, 2001), there is a dearth of research examining gender differences from the perspective of the work on expertise. This is important because the literature on expertise in science focuses almost exclusively on cognitive factors, whereas the literature explaining gender differences in science tends to focus primarily on motivation. If K-12 and college-level females are failing to achieve expertise, it makes sense to look at gender differences from the perspective of the work on expertise. What little that has been done looking at gender differences in physics problem solving and conceptual understanding has involved broad assessments, such as standardized achievement tests that do not get at the specific skills assessed by the expertise research. We believe that the expert–novice paradigm can provide a firm theoretical framework for conducting research that may help explain the well-documented gender differences in physics.

Our purpose for this study was to test a model of expertise that included gender, motivation, strategy use, categorization, and pictorial representation as predictors of physics achievement. Our expertise framework differs from what is found in the physics expertise literature in that it also considers motivation, a variable not typically identified in the physics literature, to be important for

Gita Taasoobshirazi, Department of Educational Psychology, University of Nevada, Las Vegas; Martha Carr, Department of Educational Psychology and Instructional Technology, University of Georgia.

Correspondence concerning this article should be addressed to Gita Taasoobshirazi, Department of Educational Psychology, University of Nevada, 4505 South Maryland Parkway, Las Vegas, NV 89154-3003. E-mail: gita.tshirazi@unlv.edu

the acquisition of expert performance. This study allowed us to determine which variables were most critical for achievement and which variables were influenced by gender. A review of the research on the variables included in the model is presented in detail below.

Problem-Solving Strategies

Within the domain of physics, novices tend to use the working-backward strategy when solving problems (Chi, 2006; Larkin, 1985; Williams & Noyes, 2007). Novices often work from the desired goal, breaking the problem into subgoals, each with associated equations, and continue this process until there are no unknowns left to be calculated. Then they reverse and work backwards to insert the newly found subgoal values into the preceding equations in order to solve the problem. The process is data driven with a goal of performing calculations to solve equations to find unknowns. Experts, in contrast, tend to use the working-forward strategy (Chi, 2006; Davidson, 2003; Williams & Noyes, 2007). When solving physics problems, experts work forward from a set of equations generated from the information provided in the problem, concluding the solution sequence with the goal of the problem. This process is more goal directed and cognitively efficient (i.e., exerting less cognitive load on working memory) than the working-backward strategy (Davidson, 2003; Sweller, 2003).

While the working-forward and working-backward strategies may result in the same answer, the working-forward strategy is considered to be purposeful problem solving (Snyder, 2000). In contrast, the working-backward strategy involves manipulating equations with almost no planning and little conceptual understanding of what is being done (Williams & Noyes, 2007). Without having an understanding of the equations being used and the direction the problem solver is going, a novice is likely to find him- or herself unable to solve the problem. It is for this reason that the working-forward strategy is more likely to lead to the correct solution (e.g., Zajchowski & Martin, 1993).

Although the bulk of the research indicates that the working-forward strategy is a characteristic of experts, there is research that indicates that novices can use this strategy. Priest and Lindsay (1992) examined 49 novices (university students in a college of education or a college of grammar) to 30 experts (postgraduate students in the physics department) as they solved six physics problems. They found that there was no significant difference between the novices and experts in their use of the working-forward and working-backward strategies. In fact, they found that both groups mainly used the working-forward strategy. However, these results may be due to the fact that subjects who failed to answer all six problems correctly were dropped from the study (24 novices were dropped from the study, and no experts were dropped). Because the working-forward strategy is more likely to result in the correct answer, it is likely for this reason that most subjects were found to use the working-forward strategy and no differences were found between the novices and experts.

In another study, Zajchowski and Martin (1993) examined 10 introductory-level college physics students solving mechanics problems and thinking aloud as they solved the problems. They found that the more novice problem solvers (as assessed by a pretest), who had been instructed to conceptually analyze problems

prior to solving them, were using the working-forward strategy to the same extent as the more expert problem solvers. The authors suggested these results were due to novices' practice conceptually analyzing mechanics problems. Although the research indicates that we should see more of the working-forward strategy among students in more advanced physics courses, its use is not necessarily restricted to that group. One of our goals of this study was to determine how well strategy use discriminates between more expert and novice students, and to examine the extent to which these strategies are linked to conceptual knowledge of physics.

Pictorial Representation

The ability to pictorially depict key variables and their relationships in physics problems is a major difference between expert and novice problem solvers. Before solving a problem, experts will represent the relationships in the problem by sketching a picture of the problem. Novices, in contrast, tend to focus solely on setting up and solving equations (e.g., Dhillon, 1998; Larkin, McDermott, Simon, & Simon, 1980; Larkin & Simon, 1987; Stylianou & Silver, 2004). Physics textbooks, high school and college physics instructors, and researchers emphasize the importance of pictorial representations in physics because pictorially representing a problem allows the problem solver to determine which approach to the problem is appropriate, to identify the forces and energies at work, and to reduce the amount of information that must be attended to at one time (diSessa, 2004; Friedman & diSessa, 1999; Hewitt, 2005). Further, a pictorial representation allows the problem solver to visualize the role and interaction of the various factors in a problem. Pictorially representing problems before beginning to work on calculations is particularly important as problems become more complex and additional factors (e.g., angles, forces) begin to play a role in the problems. Van Heuvelen (1991) suggested that novices often fail to draw a sketch of the problems they are solving because they do not understand the concepts and principles involved in the problems. Thus, as novices progress toward expertise and gain more conceptual knowledge, the use of pictorial representations is likely to emerge, and the quality of those representations is likely to increase.

Conceptual Knowledge and Problem Categorization

One factor that has a significant impact on expert performance is domain-specific knowledge and the organization of that knowledge. Experts in all subject domains have greater domain knowledge, and this knowledge is better organized in comparison to that of novices (Alexander, 2003; Hatano & Oura, 2003; Sabella & Redish, 2007). In physics, evidence of the difference between experts and novices in their conceptual knowledge and how they store, relate, and use this knowledge can be found in how they categorize problems (Chi, Feltovich, & Glaser, 1981; Snyder, 2000; Zimmerman & Campillo, 2003). When categorizing problems, experts focus on the principles and laws underlying the problems, while novices tend to focus on superficial (surface) features of the problems.

Studies of problem categorization indicate that expert problem solvers tend to view two problems as similar when the same law or principle can be applied to solve the problems. Novice problem solvers, in contrast, tend to view two problems as similar when the

problems share the same surface features such as terminology or objects (Chi et al., 1981; Chi & Slotta, 1993; Williams & Noyes, 2007). Differences in categorization are expected to be linked to differences in strategy use with principle-based categorizations supporting the working-forward strategy (Schneider, 1993). For the current study, we expected problem categorization to directly predict both pictorial representation and strategy use.

Motivation

Although not described in the expertise literature in physics, the role of motivation in influencing expertise is expected to be critical in any domain (e.g., Ericsson, 1996, 2004, 2006). Students who are highly motivated engage in behaviors and practice that lead to more expert knowledge and skills and, ultimately, high performance (Ericsson, 2006; Ericsson, Krampe, & Tesch-Romer, 1993). Research indicates that the important components that should be taken into account when examining students' motivation to learn science include intrinsic motivation, extrinsic motivation, task relevancy, self-determination, self-efficacy, and assessment anxiety (e.g., Duncan & McKeachie, 2005; Glynn & Koballa, 2006; Pintrich & Schunk, 2002).

Motivation to perform a task for its own sake is intrinsic, whereas motivation to perform a task as a means to an end is extrinsic (Ryan & Deci, 2000). Students who are intrinsically motivated work on a task because they find it interesting; students who are extrinsically motivated work on a task to attain a desirable outcome such as a good grade. Both types of motivation, however, are important in contributing to students' success in their courses (Ryan & Deci, 2000).

Other important components of motivation include task relevancy and self-determination. How relevant students see a task to their personal goals has been found to influence time spent on the task (Pintrich & Schunk, 2002). Self-determination refers to students having some choice in and control over their learning (Reeve, Hamm, & Nix, 2003). When college science students have the opportunity to choose what their assignments will be, they are more likely to learn from the assignments (Glynn & Koballa, 2006).

Self-efficacy refers to students' beliefs about their capabilities in a specific area, which influences choice of activities and achievement (Bong, 2001; Chemers, Hu, & Garcia, 2001). Zusho and Pintrich (2003) found that even after controlling for prior achievement, students' self-efficacy was the best predictor of grades in an introductory-level college chemistry course. Finally, assessment anxiety is an important component of motivation. A high level of assessment anxiety has been found to interfere with a student's performance on a task, and students perform best when their level of anxiety is at a low to moderate level (e.g., Cassady & Johnson, 2002).

Although the work by Ericsson (2006) focuses on the willingness to apply considerable practice to increase expertise, high levels of practice are a symptom of a number of underlying motivations. For instance, students who engage in practice both see the domain as relevant and expect that they will be able to achieve expertise in that domain. We expected to find motivation to be linked to better strategy use, categorization, pictorial representation, and achievement. We also expected our more expert group to have higher motivation.

Gender and Expertise in Physics

The largest gender differences in achievement and participation exist in the physical sciences (Katz, Allbritton, Aronis, Wilson, & Soffa, 2006), particularly in physics (Larose, Ratelle, Guay, Senecal, & Harvey, 2006; National Assessment of Educational Progress, 2005; NSF, 2004). Males from elementary school to college have been found to have higher physics achievement, have a more positive attitude toward physics, have greater physics self-efficacy, take more physics courses, and be more likely to major in physics (e.g., NSF, 2004). While the research on gender differences in science achievement has focused primarily on the poor motivation of females in K-12 and in college (e.g., Morgan et al., 2001), there is little research comparing the problem-solving strategies and conceptual understanding of males and females across these grade levels. Thus, researchers do not know whether females' poorer performance is linked to less expert strategy use and conceptual understanding.

The few studies that have examined gender differences in students' strategy use in science focused on students' learning strategies rather than on problem-solving strategies. These studies examined whether students were focusing on the conceptual aspects of the material when learning science or on the rote memorization of facts. The results of the research on middle school students is inconclusive with one study (Nolen, 1988) indicating that girls are using more conceptual strategies than boys and another study (Meece & Jones, 1996) indicating no gender differences in the use of conceptual and rote strategies. The one study examining strategy use in an older population in chemistry (Atkin, 1977) indicates gender differences in favor of college men's use of better, conceptual-based strategies. The work that has been done mainly uses self-reports to question students about general strategies for learning science, as opposed to domain-specific strategies.

In physics, there is an absence of published research comparing males to females across all grade levels in regard to the working-forward and working-backward strategies, and examining how these strategies influence physics achievement. Although there is a large body of research examining experts and novices in physics, we were unable to locate a study addressing gender differences within and across expertise categories. Females may be using less productive strategies in physics, possibly contributing to their poor achievement and participation. For instance, middle school and high school girls tend to describe science as facts to memorize (Kahle & Lakes, 1983; Taasoobshirazi & Carr, 2008), suggesting that they may use rote memorization to learn physics. For physics, rote memorization is negatively related to classroom achievement (Cavallo, Rozman, Blickenstaff, & Walker, 2004) and would likely lead to the use of the working-backward strategy. Although there is a dearth of research examining gender differences in students' problem categorizations, if females are using memorization to learn physics, their knowledge base may be poorly organized in a way that would result in categorizations based on superficial features of problems. The research on gender differences in science suggests that females would use less expert strategies and be poorer than males at problem categorization. These differences may be driven by females' well-documented low motivation in physics.

Present Study

The present study tests a model of expertise in physics that includes strategy use, pictorial representation, problem categorization, motivation, and gender as predictors of physics achievement. The model was tested on two different groups of introductory-level physics students. The first group comprised students in an introductory physics course required for science majors, while the second group comprised students enrolled in a more advanced introductory physics course required for physics and engineering majors.

Based on the existing research on expert and novice differences in physics and the research on gender differences in science, we developed the model shown in Figure 1. As can be seen in the figure, gender was expected to influence motivation, strategy use, and categorization, with men having higher motivation, using better strategies, and having better categorizations. Students with higher motivation were expected to be more inclined to engage in behaviors and practice that would support them learning the material at a deeper level, resulting in better categorizations. Similarly, students with higher motivation were expected to engage in more problem-solving practice, supporting the more expert working-forward strategy. A focus on the deeper aspects of the material and problems, as indicated by categorizations, was expected to influence student achievement both directly, and indirectly, through the use of the more expert working-forward strategy. Consistent with the research, better problem categorization was expected to support better pictorial representations. With better pictorial representations, the use of the working-forward strategy was expected to increase as a deeper understanding of the

concepts and principles and their relationship to one another increased. Finally, strategy use was expected to directly predict achievement with the working-forward strategy supporting higher achievement.

We used the structural equation modeling technique of path analysis to test the model (Kline, 2005). For both level courses, the hypothesized model was compared to two alternate models. Because the bulk of the research attempting to explain gender differences in science achievement focuses on motivation, for the first alternative model, we dropped paths from gender to strategy use and gender to categorization in order to test the fit of a model where gender influences expert performance exclusively through motivation. A second alternate model tests the hypothesis that problem categorization does not directly predict physics achievement, but influences achievement only indirectly through improved strategy use. This is because high school-level and college-level physics instruction and assignments are centered on students correctly setting up and solving problems (e.g., Henderson, 2005; McDermott, 2001). Thus, we removed the path from categorization to achievement in order to test a model where a good conceptual understanding only influences achievement through strategy use.

Method

Subjects

Subjects included a total of 374 students (185 men and 189 women) from two different-level physics courses offered at the physics departments of five different universities in Georgia. In order to recruit subjects, we e-mailed physics instructors at five universities who taught either of the two courses. These universities were selected based on proximity to the Atlanta, Georgia, area (within 100 miles). From 12 trigonometry-based courses, 8 instructors agreed to participate in the study. From five calculus-based courses, 4 instructors agreed to participate in the study. Students' participation was voluntary, following the guidelines for research with human subjects specified by the institutional review board, with informed consent forms signed by the students.

The first course was a trigonometry-based introductory-level physics course required for science majors, such as biology and chemistry majors. Students only needed high school-level algebra for the course. Eight trigonometry-based classes from four universities (two large public universities and two small public universities) participated in the study. A subtotal of 245 students (100 men and 145 women) from a total of 481 students (230 men and 251 women) enrolled in the trigonometry-based courses participated in the study. Overall, a majority of the students participated in the study (51%), and students who participated earned a small amount of extra credit.

The second physics course was a calculus-based introductory-level physics course for physics and engineering majors. The course covered the same material as the trigonometry-based course, but used calculus rather than trigonometry for much of the problem solving. Differential calculus was a prerequisite for the course. The course was a more difficult version of the trigonometry-based physics course, and students in this course were expected to be more expert at physics by the end of the semester. Four classes from three

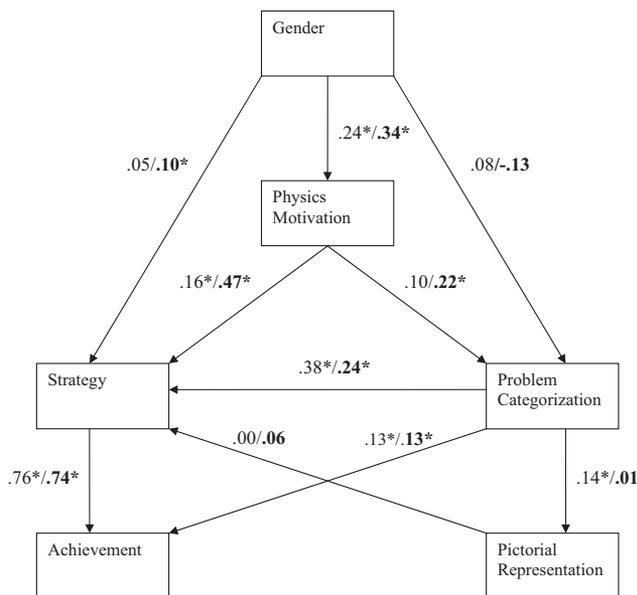


Figure 1. Theoretical model of expertise in college physics. Note that model includes path values for both the trigonometry-based physics students and the calculus-based physics students. Path values for calculus-based students are in boldface, while path values for the trigonometry-based students are unbolded. Significant values (using a cutoff value of $t = 1.96$) are marked with an asterisk.

universities (one large public university, one large private university, and a small public university) participated in the study. A subtotal of 129 students (85 men and 44 women) from a total of 181 students (120 men and 61 women) enrolled in the calculus-based courses participated in the study. A majority of the students participated in the study (71%), and students who participated earned a small amount of extra credit.

Procedure and Materials

Students in both courses were administered a packet that included a motivation questionnaire; five physics problems designed to assess strategy use, pictorial representation, and achievement; and four categorization tasks used to determine whether students were focusing on conceptual or surface features of physics problems. Information about students' gender was also collected. This packet was administered late in the fall semester after the unit on mechanics was covered. Students in both level courses spend approximately 10 of the 15 weeks of the course covering major topics in mechanics. The packet took approximately 45–60 min for students to complete. Students were allowed to use their notes and textbook when completing the packet. This was done because many physics instructors allow their students to use formula sheets, class notes, and even their books when completing physics tests. It is expected that if students do not understand the material or problems, having the use of the textbook and class notes will be of limited help. Students, however, were told to work alone when completing the packet and were on the university's academic honor code. The packets were scored by two raters: Gita Taasobshirazi, who is an educational psychologist, and a high school physics instructor. The packet can be obtained by contacting Gita Taasobshirazi.

Motivation. We assessed student motivation in physics using the Physics Motivation Questionnaire (Glynn & Koballa, 2006). The Physics Motivation Questionnaire includes 30 items that assess six important components of student motivation in physics including intrinsically motivated physics learning (e.g., "I enjoy learning physics"), extrinsically motivated physics learning (e.g., "I like to do better than other students on physics tests"), relevance of learning physics to personal goals (e.g., "The physics I learn relates to my personal goals"), self-determination for learning physics (e.g., "It is my fault if I do not understand physics"), self-efficacy in learning physics (e.g., "I believe I can master the knowledge and skills in physics courses"), and anxiety about physics assessment (e.g., "I become anxious when it is time to take a physics test"). Students responded to each of the 30 randomly ordered items on a 5-point Likert scale ranging from 1 (*never*) to 5 (*always*) from the perspective of "When learning physics . . ." The anxiety about physics assessment items were reverse scored when added to the total, so that a higher score on this component meant less anxiety.

Previous findings (Glynn & Koballa, 2006) indicate that the Physics Motivation Questionnaire is reliable as measured by coefficient alpha ($\alpha = .93$) and valid in terms of positive correlations with college students' physics grades, interest in physics careers, and number of physics courses taken. For this study, internal consistency was found to be $\alpha = .91$ for both the trigonometry-based course and the calculus-based course.

Strategy use, pictorial representation, and achievement. Five mechanics problems were administered to students (see Table 1). The five problems required a single solved mathematical solution. Students were asked to solve the problems, showing all of their work. Each of the problems is based on one of five major topics in mechanics including kinematics, forces and Newton's laws of motion, work and energy, impulse and momentum, and simple harmonic motion and elasticity (Cutnell & Johnson, 2007). The first problem deals with forces and Newton's laws of motion, the second problem deals with kinematics, the third problem involves impulse and momentum, the fourth problem involves work and energy, the fifth problem deals with simple harmonic motion and elasticity.

The first of the five problems was designed by Priest and Lindsay (1992) and was used in their own study assessing the working-forward and working-backward strategies. The other four problems were from a physics final examination used by a professor teaching both the trigonometry-based and calculus-based courses at a university that did not participate in the study. This was done to ensure that the problems were different ones than those the students had seen previously in the course. We used strategy-scoring guidelines based on those by Chi (2006) and Priest and Lindsay (1992) to score the strategies as either working forward or working backward. For the five problems, the less efficient working-backward strategy was identified as being used by a student when the first equation generated attempted to solve (impossibly) for the sought quantity, and additional equations attempted to solve for unknown quantities introduced by previous equations. Students then inserted these newly found quantities into the preceding equations in order to work backward to solve the problem. The more efficient working-forward strategy was identified as being used by a student when the first equation contained only a single unknown. Each subsequent equation provided the value of a single unknown quantity, resulting in equations that led to the sought quantity. Students received 0 points for using the working-backward strategy and 1 point for using the working-

Table 1

The Physics Problems Solved by the Students

1. A block of mass 7 kg starts sliding down a plane of length 5 m, inclined at an angle of 30 degrees to the horizontal. If the coefficient of friction between the block and the plane is 0.2, find the velocity (vt) of the block when it reaches the bottom of the plane.
2. An airplane flies horizontally with a speed of 300 m/s at an altitude of 400 m. Assume that the ground is level. What horizontal distance from a target must the pilot release a bomb so as to hit the target?
3. A 0.15 kg steel ball is dropped onto a steel plate where its speed just before impact and after impact is 4.5 m/s and 4.2 m/s, respectively. If the ball is in contact with the plate for .03 seconds, what is the magnitude of the average force (in N) applied by the plate on the ball?
4. An escalator is 30.0 meters long and slants 30 degrees relative to the horizontal. If it moves at 1.00 m/s, at what rate does it do work in lifting a 50.0 kg man from the bottom to the top of the escalator?
5. A 1.0 kg block is released from rest at the top of a frictionless incline that makes an angle of 37 degrees with the horizontal. An unknown distance down the incline from the point of release, there is a spring with $k = 200$ N/m. It is observed that the mass is brought momentarily to rest after compressing the spring 0.20 m. What distance does the mass slide from the point of release until it is brought momentarily to rest?

forward strategy. Consistent with the research scoring the working-forward and working-backward strategies, the strategies were found to clearly fall into one of the two categories so that strategy use for each problem was scored as either working forward or working backward. Points were summed across the problems and ranged from 0 to 5. The two raters scored the problems, and Cohen's kappa was found to be .90. Disagreements were settled through discussion.

Each problem was also analyzed for the quality of pictorial representations. This is the first study that has examined the quality of students' pictorial representations. Prior studies have only scored whether or not a picture was created during problem solving. We examined quality of representations because considerable emphasis in physics instruction is placed on the importance of students pictorially representing the problems they are solving (Hewitt, 2005). Thus, we expected that students who do not necessarily have a good understanding of the material would attempt to draw a sketch of the problem, including only basic factors (e.g., angles, length), whereas students who have a good understanding of the material would include those, as well as additional factors (e.g., normal force, force of gravity broken down into its horizontal and vertical components).

The pictures were scored by the two raters comparing the students' pictures to a target sketch of each problem. The target sketch was created by a college physics professor and included the necessary factors (e.g., angles, forces) needed to have a complete and inclusive pictorial representation. For the first problem, a complete sketch included the representation of three forces (normal force, frictional force, and force of gravity broken into its horizontal and vertical components) and one angle, resulting in a total of four main factors. In the second problem, a complete sketch included two important factors, the change in the vertical distance, and the horizontal velocity. In the third problem, a complete sketch included two important factors, the velocity of the object both before and after collision. In the fourth problem, a complete sketch included a total of four factors, including two forces (the normal force and the force of gravity broken into its horizontal and vertical components), length of the slant, and the angle of the slant. In the fifth problem, a complete sketch included the angle, spring, and length of the compression of the spring, for a total of three factors. For each sketch, students received 1 point for each main factor pictorially represented with a range of scores being between 0 and 15. We calculated the intraclass correlation coefficient to determine agreement among the two raters, which was found to be .90. Disagreements were settled through discussion.

In order to obtain a measure of achievement, students' responses on the five physics problems were scored 1 point for correct and 0 points for incorrect. However, if a simple calculation error occurred, full credit was still given for the problem. This occurred in only 35 of the 1,870 possible instances (1.87%). A simple calculation error was computational as opposed to conceptual, and there was evidence in the work shown that indicated that the student was able to correctly set up the problem. The two raters scored the problems, and Cohen's kappa was .96. Disagreements were settled through discussion. In order to check for rescoring consistency, 50 of the packets (25 from the trigonometry-based students and 25 from the calculus-based students) were rescored by the high school physics instructor. The second scoring session, which was performed 4 months after the first session, resulted in the same scores for all but 2 of the 250 problems.

Problem categorization. Four problem categorization tasks based on major topics in mechanics including kinematics, forces and Newton's laws of motion, work and energy, and impulse and momentum (Cutnell & Johnson, 2007) were administered to students. Each task included four physics problems, two problems from two major subtopics within a major topic. Students were not required to solve the problems but were asked to categorize the four problems into pairs and then describe in detail why they felt those particular problems went together. For the first task, concerning the topic of kinematics, the problems involved either kinematics in one dimension or kinematics in two dimensions. For the second task, forces and Newton's laws of motion, the problems involved either the equilibrium or nonequilibrium application of Newton's laws of motion. For the third task, work and energy, the problems involved either work done by a constant force or work done by a varying force. Finally for the fourth task, impulse and momentum, the problems involved either collisions in one dimension or collisions in two dimensions. The problems for the tasks came from two undergraduate trigonometry-based level physics textbooks (Cutnell & Johnson, 2007; Serway & Faughn, 2005) not being used at participating universities. We used students' categorizations and explanations to determine whether students were focusing on surface features of problems or underlying conceptual laws and principles when encountering problems. In order to deal with chance levels of correct responses to the pairings, correct explanations were needed to receive full credit for the correct pairings of problems. Specifically, for each task, students could receive up to 3 points: 1 point for correctly categorizing the problems and 1 point for each correct explanation. Students could receive up to a total of 12 points, where a higher score suggested a focus on and understanding of laws and principles when encountering physics problems. The two raters scored the explanations, and Cohen's kappa was found to be .94. Disagreements were settled through discussion.

To further check the reliability of scoring, we randomly selected and interviewed 5 students (2 from the calculus-based course and 3 from the trigonometry-based course). The students were asked about the strategies they used and about their problem categorization explanations. Cohen's kappa was found to be .95 for strategy use and .97 for categorization explanations.

Results

Described first are the results obtained from the trigonometry-based physics course. This is followed by the results obtained from the calculus-based physics course. Finally a comparison of the model variables tested in the two courses is presented. For the most part, results were similar across both courses. They differed in the more significant role gender and motivation played in the more advanced, calculus-based physics course.

Trigonometry-Based Physics

Descriptive statistics, mean comparisons, and correlations. We conducted five independent-samples *t* tests to determine whether there were any gender differences among the model variables, including motivation, pictorial representation, categorization, strategy use, and achievement. There was a significant difference and a medium effect size in the motivation ratings of the

men ($M = 98.69, SD = 14.77$) and the women ($M = 91.01, SD = 16.21$), $t(243) = 3.78, p = .00$, Cohen's $d = .50$, with the men having more motivation in physics than the women.

There was no significant difference and a negligible effect size in the pictorial representation of the men ($M = 7.01, SD = 3.42$) and the women ($M = 6.58, SD = 3.20$), $t(243) = 1.00, p = .39$, Cohen's $d = .13$. There was also no significant difference and a small effect size in the problem categorization of the men ($M = 3.99, SD = 2.99$) and the women ($M = 3.36, SD = 2.82$), $t(243) = 1.68, p = .09$, Cohen's $d = .22$. There was no significant difference and a negligible effect size in the strategy use of the men ($M = 2.20, SD = 1.36$) and the women ($M = 2.12, SD = 1.34$), $t(243) = .43, p = .66$, Cohen's $d = .06$, with both the men and women using the working-forward strategy an average of approximately two out of five possible times. Finally, there was no significant difference and a negligible effect size in the physics achievement of the men ($M = 1.73, SD = 1.38$) and the women ($M = 1.75, SD = 1.42$), $t(243) = -0.12, p = .91$, Cohen's $d = -.01$, with both the men and the women answering about two out of five items correctly. In order to account for inflation of alpha due to multiple t tests and to minimize Type I error, we conducted a Bonferroni correction. The Bonferroni correction is a safeguard against multiple tests of statistical significance on the same data falsely giving the appearance of significance. The new alpha value to use as a threshold for judging statistical significance was found to be $p = .01$. Using this alpha, we found that the gender difference in motivation was still statistically significant.

The correlations among the model variables can be seen in Table 2. Although there was not a significant correlation between gender and strategy use and gender and categorization, there was a significant correlation between gender and motivation, with men being more motivated in physics than women. There was a significant correlation between motivation and strategy use; however, the correlation between motivation and categorization was not significant. There was not a significant correlation between pictorial representation and strategy use; however, the correlation between categorization and pictorial representation was significant, indicating that a focus on the deeper aspects of problems is related to more advanced pictorial representations. Further, there were significant and substantial correlations between categorization and strategy use and categorization and achievement, indicating that a focus on the deeper aspects of problems is related to more expert

strategy use and higher achievement in physics. Finally, strategy use was significantly and highly correlated with achievement. Thus, the use of the more expert working-forward strategy is related to higher achievement. Chi-square tests on the data for the trigonometry-based physics students indicated that more problems were correctly solved using the working-forward strategy (394) than the working-backward strategy (31), and fewer problems were incorrectly solved using the working-forward strategy (148) than the working-backward strategy (652; all $ps < .001$).

Model testing. We used structural equation modeling to test the model. There were no missing data values. Before empirically testing the model, we examined the data for normality and homoscedasticity. Based on the data plots (histograms of the variables), examination of skewness and kurtosis statistics (see Table 2), and Mardia's coefficient = .93, the data met the assumptions of both univariate and multivariate normality. Based on a DeCarlo macro test, no skewness, kurtosis, or outliers were found, also suggesting normality of the data. We used LISREL Version 8.80 (Jöreskog & Sörbom, 2006a), with a covariance matrix generated by PRELIS Version 2.80 (Jöreskog & Sörbom, 2006b), to test the model by means of the maximum likelihood method of estimation. This method was used because the data were normally distributed.

The overall fit of the model was very good, as indicated by a number of fit indices, all of which are described in detail in the paragraphs below. Any given index evaluates only a particular aspect of model fit. Therefore, to evaluate the fit of the model, it is recommended that several fit indices be used (Kline, 2005). First, the chi-square statistic was used. The chi-square is a fit index that addresses the degree to which the variances and covariances implied by the model match the observed variances and covariances. A nonsignificant chi-square indicates that the model is a good representation of the underlying covariance matrix. The chi-square, $\chi^2(5) = 6.18, p = .29$, indicated a good fit because the p value was greater than .05. Further, the χ^2/df ratio was 1.24, suggesting a good fit based on Kline's (1998) rule that values less than three indicate a good fit.

The standardized root-mean-square residual (SRMR) is an index based on the residuals between the observed and estimated covariance matrices, with a value below .08 indicating a good fit (Hu & Bentler, 1999). The SRMR for this model was .03. The Steiger-Lind root-mean-square error of approximation (RMSEA) assesses a lack of fit of the population data to the estimated model. The RMSEA for this model was .03, which was below the cutoff value of .06 suggested by Hu and Bentler (1999).

The incremental fit index (IFI) is a fit index that is sensitive to model misspecification but not to sample size (Bentler, 1990; Hu & Bentler, 1999; Widaman & Thompson, 2003), making it a valuable indication of fit. A value greater than .95 is considered to indicate a good fit (Hu & Bentler, 1999). The value for this model was 1.00. Finally, the adjusted goodness-of-fit index (AGFI) is a measure of the proportion of the observed covariance that is accounted for by the model. A value greater than .90 is considered to indicate a good fit (Schumacker & Lomax, 1996). The value for this model was .97.

We used path analysis to estimate the direct and indirect effects in the model, control for the correlations among the hypothesized causal variables, and decompose the observed correlations into their component parts. The standardized path values and their associated t values for the model are reported in Table 3; the model

Table 2
Correlation Matrix, Means, Standard Deviations, Skewness, and Kurtosis for the Trigonometry-Based Physics Course

Variable	1	2	3	4	5	6
1. Gender	—					
2. Motivation	.24**	—				
3. Categorization	.12	.12	—			
4. Pictorial	.06	.13*	.14*	—		
5. Strategy	.03	.19**	.39**	.07	—	
6. Achievement	.00	.18**	.43**	.04	.81**	—
<i>M</i>		94.14	3.62	6.76	2.16	1.74
<i>SD</i>		16.06	2.90	3.29	1.35	1.40
Skewness		-.23	.71	-.39	.14	.48
Kurtosis		.64	-.47	-.56	-.86	-.71

* $p < .05$. ** $p < .01$.

Table 3
Decomposition of Effects in the Model for the Trigonometry-Based Physics Students

Predictor and criterion	Direct effect		Indirect effect	
	PC	<i>t</i>	PC	<i>t</i>
Gender				
Motivation	.24	3.78		
Strategy	.05	0.43	.08	2.52
Pictorial			.01	1.34
Categorization	.08	1.68	.02	1.41
Achievement			.04	0.66
Motivation				
Strategy	.16	3.00	.04	1.47
Pictorial			.01	1.25
Categorization	.10	1.51		
Achievement			.16	3.00
Categorization				
Strategy	.38	6.44	.00	0.06
Pictorial	.14	2.22		
Achievement	.13	7.25	.29	6.11
Strategy				
Achievement	.76	19.30		
Pictorial				
Strategy	.00	0.06		
Achievement			.00	0.06

Note. A cutoff value of $t = 1.96$ was used to determine whether paths were statistically significant. In terms of the relative size and influence of the standardized path coefficients, paths ranging from .05 to .10 are considered small but meaningful influences. Paths ranging from .11 to .25 are moderate in size and influence, and paths above .25 may be considered large in size and influence (Keith, 1993). PC = standardized path coefficient.

with standardized path values can be seen in Figure 1. We used a cutoff value of $t = 1.96$ to determine whether direct and indirect paths were statistically significant. In terms of size and influence of the standardized path values, we used Keith's (1993) recommended criteria: Standardized path values ranging from .05 to .10 are small but meaningful influences; path values ranging from .11 to .25 are moderate in size and influence, and path values above .25 are large in size and influence. The criterion R^2 (proportion of variance explained) by motivation was .06, by pictorial representation was .02, by categorization was .02, by strategy use was .18, and by achievement was .68.

Decomposition of effects. The decomposition of effects can be seen in Table 3. Of the 10 direct paths, 6 were significant. Gender had a significant and moderate influence (.24) on motivation, with men having greater motivation than women. Motivation had a significant and moderate influence (.16) on strategy use. However, the path from motivation to categorization was not significant. Thus, it appears that students who are motivated spend their time engaging in problem-solving practice that supports better strategy use. This may be because so much focus in physics is placed on correctly setting up and solving problems. Categorization had a significant and moderate influence (.14) on pictorial representation, indicating that a focus on the more conceptual aspects of the material and problems supports more complex and expert pictures. However, pictorial representation did not influence strategy use. Categorization had a significant and large influence (.38) on strategy use. Thus, a focus on the conceptual aspects of the material

and problems supports the use of the more expert working-forward strategy; the working-forward strategy had a significant and large influence (.76) on achievement. Finally, categorization had a significant and moderate influence (.13) on achievement, indicating that a focus on the conceptual aspects of the material and problems supports higher achievement.

Indirect paths and associated t values can also be seen in Table 3. Gender had a significant influence on strategy use through its influence on motivation and categorization. Further, motivation had a significant influence on achievement through its influence on strategy use and categorization. However, the path Motivation → Strategy → Achievement appeared to contribute to this influence, as the path from Motivation → Categorization → Achievement was negligible in size and thus played a smaller role in this indirect path. This suggests that higher motivation is linked to the use of the more expert working-forward strategy, which is linked to higher achievement. Again, this may be because so much focus in physics is placed on correctly setting up and solving problems. Thus, students who are motivated likely spend their time engaging in problem-solving practice that leads to better strategy use. However, there is also evidence that conceptual knowledge is critical in influencing strategy use and achievement: Problem categorization had a significant influence on achievement through its influence on strategy use, as the path from Categorization → Strategy → Achievement was significant and large in size (.29).

Thus, in addition to the direct paths indicating the significant influence of motivation and categorization on strategy use and achievement, the indirect paths Motivation → Strategy → Achievement and Categorization → Strategy → Achievement were found to be critical. These paths indicate that efforts to improve motivation and conceptual understanding would be useful in improving both strategy use and achievement.

Alternative models. The bulk of the research attempting to explain gender differences in science achievement focuses on motivation. Therefore, we tested an alternative model where gender influences expert performance exclusively through its influence on motivation. To do this, we removed the paths from Gender → Strategy and Gender → Categorization, and the model was retested. Results indicate a good model fit as indicated by $\chi^2(7) = 8.52$, $p = .29$, χ^2/df ratio = 1.22, SRMR = .04, RMSEA = .03, IFI = .99, and AGFI = .97. The direct and indirect path values stayed almost identical as far as size and identical as far as significance. We conducted a chi-square difference test to see whether there was a significant difference between the two models. The chi-square difference test was $\chi^2(2) = 2.34$. The critical value of a chi-square distribution with two degrees of freedom at $p = .05$ is 5.99. Because 2.34 is less than 5.99, it is concluded that there is no significant difference between the two models. Removing the two paths did not make a difference in the fit of the model because the direct path values were small and not significant.

The second alternate model assumes that conceptual knowledge (problem categorization) supports increased achievement only indirectly through improved strategy use. This is because high school-level and college-level physics instruction and assignments are centered on students correctly setting up and solving problems. Thus, the direct path from Categorization → Achievement was removed from the model. Results indicate an adequate model fit as indicated by $\chi^2(6) = 16.78$, $p = .01$, χ^2/df ratio = 2.80, SRMR = .04, RMSEA = .08, IFI = .96, and AGFI = .92. The direct and indirect

path values stayed almost identical as far as size and identical as far as significance. We conducted a chi-square difference test to see whether there was a significant difference between the two models. The chi-square difference test was $\chi^2(1) = 10.60$. The critical value of a chi-square distribution with one degree of freedom at $p = .05$ is 3.84. Because 10.60 is greater than 3.84, it is concluded that there is a significant difference between the two models and that the model with the path Categorization \rightarrow Achievement is a better fit model. Thus, it appears that a good conceptual understanding influences achievement beyond what is accounted for by strategy use. However, the paths from Categorization \rightarrow Strategy (.38) and Categorization \rightarrow Strategy \rightarrow Achievement (.29) are both larger than the direct path from Categorization \rightarrow Achievement (.13), emphasizing the importance of the role of conceptual understanding on strategy use, and in turn, achievement.

Thus, we conclude that for this group of students, our initial theoretical model, which is presented, in Figure 1 is the best fit model. Dropping the paths from Categorization \rightarrow Achievement resulted in a worse fit model. Although there was no difference between our initial hypothesized model and the model in which the paths from Gender \rightarrow Strategy and Gender \rightarrow Categorization were eliminated, we feel that theoretically these paths are important and support leaving them in.

Calculus-Based Physics

Descriptive statistics, mean comparisons, and correlations.

We conducted five independent-samples *t* tests to determine whether there were any gender differences among the model variables, including motivation, pictorial representation, categorization, strategy use, and achievement. There was a significant difference, and a medium effect size, in the motivation ratings of the men ($M = 104.85, SD = 15.27$) and the women ($M = 92.86, SD = 17.43$), $t(127) = 4.03, p = .00$, Cohen's $d = .71$, with the men having more motivation in physics than the women.

There was no significant difference and a negligible effect size in the pictorial representation of the men ($M = 8.40, SD = 2.87$) and the women ($M = 8.84, SD = 2.52$), $t(127) = -0.86, p = .39$, Cohen's $d = -.16$. There was also no significant difference and a negligible effect size in the problem categorization of the men ($M = 3.69, SD = 2.70$) and the women ($M = 4.02, SD = 2.90$), $t(127) = -0.64, p = .52$, Cohen's $d = -.12$. There was, however, a significant difference and a medium effect size in the strategy use of the men ($M = 3.22, SD = 1.34$) and the women ($M = 2.55, SD = 1.23$), $t(127) = 2.80, p = .00$, Cohen's $d = .52$, with the men using the working-forward strategy almost 15% more than the women. Finally, there was a significant difference and a medium effect size between the physics achievement of the men ($M = 2.68, SD = 1.45$) and the women ($M = 1.95, SD = 1.22$), $t(127) = 2.85, p = .00$, Cohen's $d = .54$, with the men, on average, answering almost one more problem correctly. Once again, a Bonferroni correction was conducted, providing a new alpha value of $p = .01$. Using this alpha, statistically significant gender differences were found again in motivation, strategy use, and achievement.

The correlations among the model variables can be seen in Table 4. Although there was a significant correlation between gender and strategy use and gender and motivation, the correlation between gender and categorization was not significant. There was a significant correlation between motivation and strategy use and

Table 4
Correlation Matrix, Means, Standard Deviations, Skewness, and Kurtosis for the Calculus-Based Physics Course

Variable	1	2	3	4	5	6
1. Gender	—					
2. Motivation	.34**	—				
3. Categorization	-.06	.18*	—			
4. Pictorial	-.08	.06	.01	—		
5. Strategy	.24**	.55**	.31**	.08	—	
6. Achievement	.25**	.46**	.37**	.00	.78**	—
<i>M</i>		100.76	3.81	8.26	2.99	2.43
<i>SD</i>		16.96	2.76	2.73	1.34	1.41
Skewness		-.26	.77	-.52	-.30	-.03
Kurtosis		.04	.14	.37	-.65	-.79

* $p < .05$. ** $p < .01$.

a smaller but still significant correlation between motivation and categorization. This indicates that higher motivation is related to better strategy use, but also categorization for this group. There was not a significant correlation between pictorial representation and strategy use or between categorization and pictorial representation. There were significant and substantial correlations between categorization and strategy use and categorization and achievement, indicating that a focus on the conceptual aspects of the material and problems is related to more expert strategy use and higher achievement. Finally, strategy use was significantly and highly correlated with achievement, indicating that the use of the more expert working-forward strategy is related to higher achievement. Chi-square tests on the data for the calculus-based physics students indicated that more problems were correctly solved using the working-forward strategy (278) than the working-backward strategy (26), and fewer problems were incorrectly solved using the working-forward strategy (106) than the working-backward strategy (235; all $ps < .001$).

Model testing. We used structural equation modeling to test the model. There were no missing data values. Before empirically testing the model, we examined the data for normality and homoscedasticity. Based on the data plots (histograms of the variables), examination of skewness and kurtosis statistics (see Table 4), and Mardia's coefficient = .96, the data met the assumptions of both univariate and multivariate normality. Based on a DeCarlo macro test, no skewness, kurtosis, or outliers were found, also suggesting normality of the data. We tested the model by means of the maximum likelihood method of estimation because the data were normally distributed.

The fit of the model was very good. The $\chi^2(5) = 4.82, p = .44$, χ^2/df ratio = .96, SRMR = .03, RMSEA = .00, IFI = 1.00, and AGFI = .95. Figure 1 includes the model with standardized path values. The standardized path values and their associated *t* values for the model are reported in Table 5. We used a cutoff value of $t = 1.96$ to determine whether direct and indirect paths were statistically significant. The paths are described using Keith's (1993) suggested criterion as previously explained. The criterion R^2 (proportion of variance explained) by motivation was .11, by pictorial representation was .00, by categorization was .05, by strategy use was .36, and by achievement was .63.

Decomposition of effects. The decomposition of effects can be seen in Table 5. Of the 10 direct paths, seven were significant.

Table 5
Decomposition of Effects in the Model for the Calculus-Based Physics Students

Predictor and criterion	Direct effect		Indirect effect	
	PC	<i>t</i>	PC	<i>t</i>
Gender				
Motivation	.34	4.02		
Strategy	.10	2.88	.15	2.63
Pictorial			.00	-0.16
Categorization	-.13	-0.63	.07	2.06
Achievement			.18	2.50
Motivation				
Strategy	.47	6.72	.05	1.93
Pictorial			.00	0.16
Categorization	.22	2.40		
Achievement			.42	6.10
Categorization				
Strategy	.24	3.26	.00	0.16
Pictorial	.01	0.16		
Achievement	.13	4.03	.18	3.16
Strategy				
Achievement	.74	12.99		
Pictorial				
Strategy	.06	0.87		
Achievement			.04	0.87

Note. A cutoff value of $t = 1.96$ was used to determine whether paths were statistically significant. In terms of the relative size and influence of the standardized path coefficients, paths ranging from .05 to .10 are considered small but meaningful influences. Paths ranging from .11 to .25 are moderate in size and influence, and paths above .25 may be considered large in size and influence (Keith, 1993). PC = standardized path coefficient.

Gender had a significant and large influence (.34) on motivation, with the men having greater motivation than the women. Further, gender had a significant and small influence (.10) on strategy use for this group. Thus, the men appeared to be more likely to use the working-forward strategy. Student motivation had a significant and large influence (.47) on strategy use. Thus, higher motivation supports greater use of the more expert working-forward strategy. The influence of motivation on categorization was significant and moderate (.22), indicating that higher motivation encourages students to focus on the deeper aspects of the material and problems. Thus, it appears that for this higher level course that students who were motivated focused on attaining both good problem-solving skills and a good conceptual understanding. Categorization had a significant and moderate influence on strategy use (.24) and achievement (.13). Thus, a focus on the deeper aspects of the material and problems supports the use of the more expert working-forward strategy and higher achievement; the working-forward strategy also had a significant and large (.74) influence on achievement. Pictorial representation did not play a role in the model. Thus, for the calculus-based students, two additional paths were found to be significant: the path from gender to strategy use and the path from motivation to categorization.

Indirect paths and associated t values can be seen in Table 5. Gender had a significant influence on strategy use through its influence on motivation and categorization. However, because the path from Gender → Categorization → Strategy was negligible in size, the path from Gender → Motivation → Strategy played a

larger role in this influence, indicating that men have higher motivation, and this higher motivation supports more expert strategy use. Gender also had significant and small influence (.07) on categorization through its influence on motivation. Gender had a significant influence on achievement through its influence on strategy use and categorization, but it was the path Gender → Strategy → Achievement that contributed to this influence, as the path from Gender → Categorization → Achievement was negligible in size.

The small sized path (.05) from motivation to strategy use as mediated by categorization approached significance ($t = 1.93$), indicating that high motivation supports a focus on the more conceptual aspects of problems, which in turn supports the use of the more expert working-forward strategy. Motivation had a significant and large influence (.42) on achievement through its influence on strategy use and categorization. However, it was the path from Motivation → Strategy → Achievement that appeared to contribute to this influence, as the path from Motivation → Categorization → Achievement was negligible in size. Finally, the path from categorization to achievement as mediated by strategy use was moderate (.18) and significant, suggesting that a focus on the conceptual aspects of the material and problems supports the use of the more expert working-forward strategy, which supports higher achievement.

In addition to the direct paths, the indirect paths suggest that working to improve motivation and conceptual understanding will be useful in improving strategy use and achievement. These results are analogous to those found for the trigonometry-based course. However, gender played a larger role for the calculus-based students, as seen in both the direct and indirect paths.

Alternative models. The alternative model where gender influences expert performance exclusively through its influence on motivation was tested. Thus, the paths from Gender → Strategy and Gender → Categorization were removed. Results indicate a good model fit as indicated by $\chi^2(7) = 8.59, p = .28, \chi^2/df$ ratio = 1.23, SRMR = .04, RMSEA = .04, IFI = .99, and AGFI = .94. The direct and indirect path values stayed almost identical as far as size and identical as far as significance. A chi-square difference test indicated that removing the two paths did not make a difference in the fit of the model. This suggests that motivation is a critical variable in influencing gender differences in expert performance.

The second alternate model assumes that conceptual knowledge (problem categorization) leads to increased achievement only indirectly through improved strategy use. Thus, the direct path from Categorization → Achievement was removed from the model. Results indicate a good model fit as indicated by $\chi^2(6) = 10.17, p = .12, \chi^2/df$ ratio = 1.70, SRMR = .03, RMSEA = .07, IFI = .98, and AGFI = .91. The direct and indirect path values stayed almost identical as far as size and identical as far as significance. A chi-square difference test indicated that the model with the path from Categorization → Achievement is a better fit model. However, the paths from Categorization → Strategy (.24) and from Categorization → Strategy → Achievement (.18) are both larger than the direct path from Categorization → Achievement (.13), emphasizing the importance of the role of a good conceptual understanding on strategy use and, in turn, achievement.

Just like for the trigonometry-based students, we conclude that for this group of students, our initial theoretical model is the best

fit model. Dropping the paths from Categorization \rightarrow Achievement resulted in a worse fit model. Although there was no difference between our initial hypothesized model and the model in which the paths from Gender \rightarrow Strategy and Gender \rightarrow Categorization were eliminated, we feel that theoretically these paths are important and support leaving them in.

Comparison of Trigonometry-Based Physics and Calculus-Based Physics

In order to determine whether there were differences between the two courses in the variables that influence expertise in physics, we conducted five independent-samples *t* tests. Results indicate that there was a significant difference, with a small effect size, in the motivation of the trigonometry-based physics students ($M = 94.14$, $SD = 16.06$) and calculus-based students ($M = 100.76$, $SD = 16.96$), $t(372) = 3.72$, $p = .00$, Cohen's $d = .40$, with the calculus-based students having significantly more motivation than the trigonometry-based students.

There was not a significant difference and a negligible effect size in the problem categorization of the trigonometry-based students ($M = 3.62$, $SD = 2.90$) and the calculus-based students ($M = 3.81$, $SD = 2.76$), $t(372) = 0.61$, $p = .54$, Cohen's $d = .07$. There was, however, a significant difference and medium effect size in the pictorial representation of the trigonometry-based students ($M = 6.76$, $SD = 3.30$) and the calculus-based students ($M = 8.55$, $SD = 2.76$), $t(372) = 5.29$, $p = .00$, Cohen's $d = .59$, with the calculus-based students drawing more complex pictorial representations. There was also a significant difference and medium effect size in the strategy use of the trigonometry-based students ($M = 2.16$, $SD = 1.35$) and the calculus-based students ($M = 2.99$, $SD = 1.34$), $t(372) = 5.73$, $p = .00$, Cohen's $d = .62$, with the calculus-based students using the working-forward strategy almost 20% more than trigonometry-based students. Finally, there was a significant difference and almost medium effect size in the physics achievement of the trigonometry-based students ($M = 1.74$, $SD = 1.40$) and the calculus-based students ($M = 2.43$, $SD = 1.41$), $t(372) = 4.52$, $p = .00$, Cohen's $d = .49$, with the calculus-based students answering almost one more item correctly. Once again, a Bonferroni correction was conducted, providing a new alpha value of $p = .01$. Using this alpha, we once again found statistically significant differences in motivation, pictorial representation, strategy use, and achievement.

Discussion

Our purpose for this study was to develop and test a theoretical model of expert performance in physics across two levels of expertise. This is the first study in which researchers tested a model in which the influences of motivation, categorization, pictorial representation, and strategy use on achievement were simultaneously tested. The current study provided an opportunity for us to determine the contributions of individual variables when other variables were considered. We also addressed the role of motivation in expert performance in physics. Finally, we examined the role of gender in influencing expert performance.

Results indicate that the model did a good job explaining the relationship among the variables for both levels of students. For both groups, motivation had a significant influence on students'

strategy use (and categorization for the calculus-based model). This motivation was greater among men than women. Categorization had a significant influence on strategy use and achievement. Strategy use had a significant influence on achievement, and men in the calculus-based course were more likely than the women to use the working-forward strategy. Pictorial representation played little role in the model for either group. For both groups of students, results indicate important indirect paths. Motivation had a significant impact on achievement as mediated by strategy use and categorization, particularly strategy use, and categorization had a significant influence on achievement as mediated by strategy use.

Strategy use played a central role in the model, influencing problem-solving success more than any other model component. We found that the students who used a working-forward strategy tended to answer the problems correctly, whereas those who used a working-backward strategy tended to answer the problems incorrectly. While there is evidence that students in introductory-level physics courses can be taught the working-forward strategy (Zajchowski & Martin, 1993), our data indicate that this strategy is somewhat linked to students' understanding of the principles that underlie the solution to that problem. We suggest that when teaching students the working-forward strategy that instructors concurrently attend to students' conceptual knowledge. The working-forward strategy is a tool that enables students to efficiently apply their knowledge in order to engage in purposeful and goal-directed problem solving (Snyder, 2000). Thus, instructors should ensure that students have the conceptual knowledge needed to solve physics problems when teaching the working-forward strategy and should ensure that students activate their conceptual knowledge when solving problems.

Results of this study indicate that conceptual knowledge has a significant impact on problem solving and achievement. The current findings support the small number of studies pointing to the need for instruction for conceptual understanding in physics (e.g., Taconis, Ferguson-Hessler, & Broekkamp, 2001). In practice, however, most physics problems assigned for homework, administered on assessments, and presented at both the high school and college level focus students on manipulating equations to solve for a single unknown quantity, and little emphasis is placed on the underlying concepts (Briscoe & Prayaga, 2004; Kang & Wallace, 2005). Instructors should tailor their curriculum and instruction so that a strong conceptual understanding of the material is stressed.

An excellent way to improve problem solving while also attending to conceptual knowledge is through the use of physics instructional software programs. One popular program is the Andes Physics Tutoring System, which was developed by the University of Pittsburgh and the United States Naval Academy (VanLehn et al., 2005). The Andes system provides detailed feedback to students as they solve introductory-level physics problems. At each step of the problem-solving process, Andes system tells students whether their entry is correct or incorrect and provides remedial responses, in the form of principle-based hints, to help support the students' problem solving. Semester-long evaluations of the program showed that students in physics classes who used the Andes system had higher problem-solving success than did students who did not use the program (VanLehn et al., 2005).

The Andes system includes over 500 problems that cover most topics in standard two-semester introductory-level physics courses.

The program and supplementary materials can be downloaded at no cost because the Andes system is funded by the Cognitive Science Program of the Office of Naval Research and the National Science Foundation. The program is designed to be used by introductory-level high school and college physics students.

The results of this study replicate prior research (e.g., Murphy & Whitelegg, 2006) indicating substantial motivational differences between K-12 and college-level males and females in physics. It extends prior research by showing that gender influences strategy use and problem categorization indirectly through motivation. It also shows that, for the more advanced course, women were less likely to use the more advanced working-forward strategy, and this was tied to poor motivation. There is a clear need to improve the motivation of women in physics, but there is also a clear need to foster the development of more advanced strategies. Given that women and men had similar scores on problem categorization, efforts to improve motivation and strategy use should be helpful in improving women's achievement.

The results of this study indicate that efforts made to increase student motivation will be helpful for improving strategy use, conceptual understanding, and ultimately achievement, particularly for women. Gender differences in expert performance are linked to motivation. This is evident in the alternative model in which the paths from gender to strategy use and gender to categorization were removed, and only the path from gender to motivation was included. Removing the two paths made no difference in the fit of the model, indicating that motivation was the critical variable in leading to differences in expert performance. Glynn, Taasobshirazi, and Brickman (2007) found that making science more relevant to students increases their motivation, and they recommended the use of case studies that make science more relevant to students' interests, majors, and future career goals. This would be particularly beneficial for women as women, more than men, feel that physics is irrelevant to them and to their future goals (Murphy & Whitelegg, 2006).

Pictorial representation played little role in either of the models. This is surprising and informative because in both high school and college physics classrooms, there is a great deal of emphasis placed on pictorial representations. Textbooks include figures to go along with sample problems and continuously encourage students to draw free-body diagrams to represent the problems they are solving (e.g., Cutnell & Johnson, 2007; Serway & Faughn, 2005). Furthermore, high school and college instructors encourage their physics students to pictorially represent the problems they are solving and even do so in their own class examples (T. Heil, personal communication, October 30, 2006; W. Snow, personal communication, October 21, 2007). In our study, pictorial representation did not play a significant role in the calculus-based model, and although categorization did have a moderate influence on pictorial representation in the trigonometry-based model, pictorial representation did not influence strategy use. Our data indicate that even when students do a good job creating pictures, the pictures are not necessarily helpful. Given the emphasis that physics textbooks and instructors place on pictorial representations, it may be that students simply drew the pictures without using them to reason when solving the problems. Additional research needs to be conducted examining how pictorial representations are being used by students when they are solving physics problems.

Limitations and Future Research

Although not described in the expertise literature in physics, effective metacognition is considered important for efficient problem solving and for the transition from novice to expert (Feltovich, Prietula, & Ericsson, 2006; Zimmerman, 2006). Metacognition was not assessed in the current study because of the need for individual interviews during problem solving (e.g., Carr & Jessup, 1997). Future researchers should examine students' think-alouds as they solve physics problems and categorization tasks in order to understand the role of metacognition on students' problem-solving strategies and categorizations. In addition, the impact of social support is expected to be critical for the transition to expertise (Bloom, 1985; Ericsson, 2006) and should also be included in the model. Both of these variables could influence expert performance and uniquely contribute to the model's potential for explaining the variance in students' physics achievement.

Although over half of the students (51% of the students in the trigonometry-based course and 71% of the students in the calculus-based course) in each level course participated in the study, consistent with the requirements of the institutional review board for human subjects, students were not required to participate and did so by choice to earn a small amount of extra credit in the course. Thus, the students who participated may be different in their motivation and achievement levels from those who chose not to participate. We believe, however, that given the high rate of participation, the model would hold true for all students with the same relationships holding true when both lower performing students and higher performing students are included in the sample.

We used students' solution accuracy as a measure of achievement. This is in line with prior research in physics (e.g., Zajchowski & Martin, 1993). However, future research should also include a better measure of physics achievement. Although the five problems administered to students were from five major topics in mechanics, a student's grade in the course up to the end of the unit on mechanics could have provided a better measure of achievement. Unfortunately, for this study, we were unable to attain access to these grades but suggest that future researchers include a more effective measure of achievement.

From both sections of the course, some of the students who participated in the study were from minority groups, including Asian/Pacific Islander, African American, or Hispanic or Latino. However, because the percentages in each group and overall were so small, we did not treat minority status as a variable to avoid misleading statistical inferences. A larger sample size may provide the opportunity to study minority students, helping to explain the differences in achievement and participation that exist among minority students in physics (NSF, 2004).

Results of this study have important implications for instruction in these level physics courses. Given the decline in achievement and participation by all students in physics (Murphy & Whitelegg, 2006), instructors need to be doing a better job teaching physics. The results of the current study provide a description of the relationship between the variables that underlie expertise in physics and indicate which are most critical for expert performance. Results indicate that physics instructors should spend more time focusing students on the conceptual aspects of the material they are learning and help foster the motivation that supports conceptual understanding and expert strategy use. Efforts to improve the

motivation of women are particularly important. Instructors should also show students the difference between the working-forward and working-backward strategies, show them the connection between the two strategies and conceptual understanding, and encourage students to work forward when solving problems. The use of pictorial representations in physics classrooms need to be further examined, but our data indicate that the significant time and energy placed on pictorial representations of problems may be better spent working to improve students' conceptual understanding, motivation, and strategy use.

References

- Alexander, P. A. (2003). Can we get there from here? *Educational Researcher*, 32, 3–4.
- Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Anders & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 64–92). New York: Cambridge University Press.
- Atkin, J. A. (1977). An information processing model of learning and problem solving. (Doctoral dissertation, Cornell University). *Dissertation Abstracts International*, 38, 6488.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246.
- Bloom, B. S. (1985). *Developing talent in young people*. New York: Ballantine Books.
- Bong, M. (2001). Between and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, 93, 23–34.
- Briscoe, C., & Prayaga, C. S. (2004). Teaching future K-8 teachers the language of Newton: A case study of collaboration and change in university physics teaching. *Science Education*, 88, 947–969.
- Carr, M., & Jessup, D. L. (1997). Gender differences in first grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, 89, 318–328.
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27, 270–295.
- Cavallo, A. M. L., Rozman, M., Blickenstaff, J., & Walker, N. (2004). Students' learning approaches, reasoning abilities, motivational goals, and epistemological beliefs in differing science courses. *Journal of College Science Teaching*, 33, 18–24.
- Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology*, 93, 55–64.
- Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In N. Charness, P. J. Feltovich, R. R. Hoffman, & K. A. Ericsson (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). New York: Cambridge University Press.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Chi, M. T. H., & Slotta, J. D. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction*, 10, 249–260.
- Cutnell, J. D., & Johnson, K. W. (2007). *Physics*. New York: Wiley.
- Davidson, J. E. (2003). Insights about insightful problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 149–175). New York: Cambridge University Press.
- Dhillon, A. S. (1998). Individual differences within problem-solving strategies used in physics. *Science Education*, 82, 379–405.
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22, 293–331.
- Duncan, T., & McKeachie, W. J. (2005). The making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologist*, 40, 117–128.
- Ericsson, K. A. (1996). The acquisition of expert performance. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts, sciences, sports, and games* (pp. 1–50). Mahwah, NJ: Erlbaum.
- Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, 79, 70–81.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge, England: Cambridge University Press.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.
- Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In N. Charness, P. J. Feltovich, R. R. Hoffman, & K. A. Ericsson (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–67). New York: Cambridge University Press.
- Friedman, J. S., & diSessa, A. A. (1999). What students should know about technology: The case of scientific visualization. *Journal of Science Education and Technology*, 8, 175–195.
- Glynn, S. M., & Koballa, T. R. (2006). Motivation to learn college science. In J. Mintzes & W. H. Leonard (Eds.), *Handbook of college science teaching*. Arlington, VA: National Science Teachers Association Press.
- Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience majors learning science: A theoretical model of motivation. *Journal of Research in Science Teaching*, 44, 1088–1107.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities*. Mahwah, NJ: Erlbaum.
- Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational Researcher*, 32, 26–29.
- Hausler, P., & Hoffman, L. (2002). An intervention study to enhance girls' interest, self-concept, and achievement in physics classes. *Journal of Research in Science Teaching*, 39, 870–888.
- Henderson, C. (2005). The challenges of instructional change under the best of circumstances: A case study of one college physics instructor. *American Journal of Physics*, 73, 778–786.
- Hewitt, P. G. (2005). *Conceptual physics*. New York: Addison Wesley.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional fit criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Institute of Education Sciences. (2007). *Encouraging girls in math and science*. Washington DC: National Center for Education Research.
- Jöreskog, K. G., & Sörbom, D. (2006a). LISREL 8.80 [Computer software]. Chicago, IL: Scientific Software International.
- Jöreskog, K. G., & Sörbom, D. (2006b). *PRELIS 2: User's reference guide. A program For multivariate data screening and data summarization: A preprocessor for LISREL* [Computer software]. Chicago, IL: Scientific Software International.
- Kahle, J. B., & Lakes, M. K. (1983). The myth of equality in science classrooms. *Journal of Research in Science Teaching*, 20, 131–140.
- Kang, N., & Wallace, C. S. (2005). Secondary science teachers' use of laboratory activities: Linking epistemological beliefs, goals, and practices. *Science Education*, 89, 140–165.
- Katz, S., Allbritton, D., Aronis, J., Wilson, C., & Soffa, M. L. (2006). Gender, achievement, and persistence in an undergraduate computer science program. *The Database for Advances in Information Systems*, 37, 42–57.
- Keith, T. Z. (1993). Causal influences on school learning. In H. J. Walberg

- (Ed.), *Analytic methods for educational productivity* (pp. 21–47). Greenwich, CT: JAI Press.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Larkin, J. H. (1985). Understanding problem representations and skill in physics. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Vol. 2. Research and open questions* (pp. 141–159). Hillsdale, NJ: Erlbaum.
- Larkin, J. H., McDermott, J., Simon, H., & Simon, D. (1980, June 20). Expert and novice performance in solving physics problems. *Science*, 208, 1335–1342.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.
- Larose, S., Ratelle, C. F., Guay, F., Senecal, C., & Harvey, M. (2006). Trajectories of science self-efficacy and academic and vocational adjustment in science and technology programs. *Educational Research and Evaluation*, 12, 373–393.
- McDermott, L. C. (2001). Oersted Medal Lecture 2001: Physics education research—The key to student learning. *American Journal of Physics*, 69, 1127–1137.
- Meece, J. L., & Jones, G. M. (1996). Gender differences in motivation and strategy use in science: Are girls rote learners? *Journal of Research in Science Teaching*, 33, 393–406.
- Morgan, C., Isaac, J. D., & Sansone, C. (2001). The role of interest in understanding the career choices of female and male college students. *Sex Roles*, 44, 295–320.
- Murphy, P., & Whitelegg, E. (2006). *Girls in the physics classroom: A review of the research into the participation of girls in physics*. Retrieved August 5, 2007, from http://www.iop.org/activity/education/Making_a_Difference/Policy/file_6574.pdf
- National Assessment of Educational Progress. (2005). *Results of the 2005 Science Trial Urban District Assessment*. Washington, DC: National Center for Education Statistics.
- National Science Foundation. (2004). *Women, minorities, and persons with disabilities in science and engineering: 2004*. Retrieved October 5, 2007, from <http://www.nsf.gov/statistics/wmpd/pdf/nsf04317.pdf>
- Nolen, S. B. (1988). Reasons for studying: Motivational orientations and study strategies. *Cognition and Instruction*, 5, 269–287.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. *The British Journal of Psychology*, 83, 389–405.
- Reeve, J., Hamm, D., & Nix, G. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology*, 95, 375–392.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.
- Sabella, M. S., & Redish, E. F. (2007). Knowledge organization and activation in physics problem solving. *American Journal of Physics*, 75, 1017–1029.
- Schneider, W. (1993). Acquiring expertise: Determinants of exceptional performance. In K. A. Heller, F. J. Monks, & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 311–324). London: Pergamon Press.
- Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.
- Serway, R. A., & Faughn, J. S. (2005). *College physics*. Belmont, CA: Thomson Press.
- Snyder, J. L. (2000). An investigation of the knowledge structures of experts, intermediates and novices in physics. *International Journal of Science Education*, 22, 979–992.
- Stylianou, D. A., & Silver, E. A. (2004). The role of visual representations in advanced mathematical problem solving: An examination of expert–novice similarities and differences. *Mathematical Thinking and Learning*, 6, 353–387.
- Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 215–266). San Diego, CA: Academic Press.
- Taasobshirazi, G., & Carr, M. (2008). Gender differences in science: An expertise perspective. *Educational Psychology Review*, 20, 149–169.
- Taconis, R., Ferguson-Hessler, M. G. M., & Broekkamp, H. (2001). Teaching science problem solving: An overview of experimental work. *Journal of Research in Science Teaching*, 38, 442–468.
- Van Heuvelen, A. (1991). Overview, case-study physics. *American Journal of Physics*, 59, 898–907.
- VanLehn, K., Lynch, C., Schulze, K., Shaprio, J. A., Shelby, R., Taylor, L., et al. (2005). The Andes Physics Tutoring System: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15, 1–47.
- Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods*, 8, 16–37.
- Williams, D. J., & Noyes, J. M. (2007). Effect of experience and mode of presentation on problem solving. *Computers in Human Behavior*, 23, 258–274.
- Zajchowski, R., & Martin, J. (1993). Differences in the problem solving of stronger and weaker novices in physics: Knowledge strategies, or knowledge structure. *Journal of Research in Science Teaching*, 30, 459–470.
- Zimmerman, B. J. (2006). Development and adaptation of expertise: The role of self-regulatory processes and beliefs. In N. Charness, P. J. Feltovich, R. R. Hoffman, & K. A. Ericsson (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 705–722). New York: Cambridge University Press.
- Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 233–262). New York: Cambridge University Press.
- Zusho, A., & Pintrich, P. R. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education*, 25, 1081–1094.

Received January 30, 2008

Revision received November 3, 2008

Accepted November 3, 2008 ■