

A Synthesis of Discipline-Based Education Research in Physics

Jennifer L. Docktor & José P. Mestre

University of Illinois at Urbana-Champaign

| | |
|--|-----|
| Introduction | 2 |
| Conceptual Change | 4 |
| Misconceptions | 11 |
| Knowledge in Pieces or Resources | 12 |
| Ontological Categories | 12 |
| Problem Solving | 19 |
| Expert-Novice Studies | 25 |
| Worked Examples | 26 |
| Representations | 28 |
| Mathematics in Physics | 29 |
| Instructional Strategies | 31 |
| Curriculum & Instruction in Physics | 44 |
| Lecture-based Methods | 49 |
| Recitation or Discussion Methods | 52 |
| Laboratory Methods | 53 |
| Structural Changes to the Classroom Environment | 56 |
| General Instructional Strategies and Materials | 59 |
| Assessment | 78 |
| Development and Validation of Concept Inventories | 82 |
| Comparing Scores Across Multiple Measures | 86 |
| Comparing Scores Across Multiple Populations | 87 |
| Course Exams and Homework | 89 |
| Rubrics for Process Assessment | 91 |
| Complex Models of Student Learning | 92 |
| Cognition | 103 |
| Knowledge and Memory | 109 |
| Attention | 110 |
| Reasoning and Problem Solving | 112 |
| Learning | 113 |
| Attitudes and Beliefs about Learning and Teaching | 122 |
| Student Attitudes and Beliefs about Learning Physics | 126 |
| Faculty Beliefs and Values about Teaching and Learning | 132 |
| Instructor Implementations of Reformed Curricula | 134 |
| Teaching Assistants | 136 |
| Summary & Conclusions | 141 |
| Summary of Research in Physics Education from a Historical Perspective | 141 |
| What Distinguishes PER from Other DBER Fields? | 143 |
| Future Directions in Physics Education Research | 144 |

Introduction

This paper synthesizes physics education research (PER) at the undergraduate level, and was commissioned by the National Research Council to inform a study on the Status, Contributions, and Future Directions of Discipline Based Education Research (DBER)—a comprehensive examination of the research on learning and teaching in physics and astronomy, the biological sciences, chemistry, engineering, and the geosciences at the undergraduate level. PER is a relatively new field that is about 40 years old, yet it is relatively more mature than its sister fields in biology, chemistry, engineering, astronomy, and geosciences education research. Although much is known about physics teaching and learning, much remains to be learned. This paper discusses some of what the PER field has come to understand about learners, learning, and instruction in six general topical areas described below.

Topical Areas Covered and Organization

Given the breadth and scope of PER to date, we organize our synthesis around six topical areas that capture most of the past research in physics education: conceptual change, problem solving, curriculum and instruction, assessment, cognition, and attitudes and beliefs about learning and teaching. To ensure consistency in the presentation and to aid the DBER Committee in its charge, each of the six topical areas was organized under the following sections: research questions; theoretical framework; methodology, data collection/sources and data analysis; findings; strengths and limitations; areas for future studies; references. Unlike traditional scholarly papers where all references are included at the end, we included references at the end of each of the six sections for convenience. The topical areas and organization were arrived at through consensus discussions between the two authors and in consultation with the DBER study director and Committee chair. Although we did not place any restrictions on the dates of the research studies covered, the great majority of the studies cited are within the last 20 years, and the great majority of studies cited were done in the US. In addition to the six topical areas, we provide a Summary & Conclusions section at the very end.

Equally important to stating what this paper is covering is stating what has been left out, not by our choice but rather because of the Committee's focus and needs. We did not include the following: pre-college physics education research, research related to physics teacher preparation

or physics teacher curriculum, or “how to” articles describing research analysis (e.g., ways of analyzing video interviews of students).

The coverage herein is meant to be representative, not exhaustive. We apologize if we have left out a particular researcher’s favorite work—if we did, it was not intentional.

Possible Uses of this Synthesis

We envision multiple uses of this synthesis. First, it is meant to help the DBER Committee in its charge and deliberations. Second, it serves as a historical account of the field, taking stock in where PER has been and where it currently is. In addition, it provides a perspective of the status of other discipline-based education research relative to physics. Finally, the synthesis is an excellent resource for graduate students in PER, for graduate students or physics instructors in traditional physics areas who are interested in an overview of the PER field, and indeed to practicing PER faculty who may want to use it as a reference guide or as a resource in teaching graduate seminar on PER.

Conceptual Change

Conceptual development and conceptual change are among the most widely studied areas in physics education research. Starting in the 1970s, as researchers and instructors became increasingly aware of the difficulties students had in grasping fairly fundamental concepts (e.g., that contact forces do not exert forces at a distance; that interacting bodies exert equal and opposite forces on each other), investigations into the cause of those difficulties became common. Over time these conceptual difficulties have been given several labels, including misconceptions, naïve conceptions, and alternative conceptions. In this review we have chosen to use the term *misconceptions* but acknowledge that some researchers may have a preference for other terms.

Some would argue that misconceptions research marked the beginning of modern day physics education research. Early work consisted of identifying and cataloguing student misconceptions (Clement, 1982; McDermott, 1984), and there were entire conferences devoted to student misconceptions in the STEM disciplines (Helm & Novak, 1983; Novak, 1987) with thick proceedings emerging from them.

Research Questions

The research questions investigated under the conceptual change generally fall into the following categories:

Cataloguing misconceptions *What learning difficulties do students possess that are conceptual in nature? What are the most common misconceptions that interfere with the learning of scientific concepts?* Much work has gone into cataloguing preconceptions that students bring into physics classes prior to instruction, and identifying which of those are misconceptions that are in conflict with current scientific concepts. Although many of these studies address topics in mechanics (e.g. kinematics and dynamics), there have also been many studies conducted in electricity and magnetism, light and optics, thermal physics, and a few in modern physics. For a list of approximately 115 studies related to misconceptions in physics, see McDermott and Redish, 1999.

In addition, many investigations explore whether or not misconceptions persist following instruction, which in turn provide insights into the type of instruction that impacts students'

conceptual understanding. This body of work has generated numerous carefully documented studies (Ambrose, Heron, Vokos & McDermott, 1999; Brown & Clement, 1989; Goldberg & McDermott, 1986, 1987; McDermott & Shaffer, 1992) of the role of misconceptions in students' reasoning and learning, as well as several inventories to assess conceptual understanding in several physics domains that will be discussed further in the *Assessment* section (Beichner, 1994; Hestenes & Wells, 1992; Hestenes, Wells & Swackhammer, 1992; Maloney, O'Kuma, Hieggelke & VanHeuvelen, 2001; Maloney & Hieggelke, 2000; Thornton & Sokoloff, 1998).

Describing the architecture of conceptual structure. *What is the nature of scientific concepts in memory? What changes when conceptual change takes place?* Another thrust in conceptual change research attempts to describe the cognitive architecture of conceptual knowledge (i.e., how conceptual knowledge is structured in memory). This body of work has generated lively debate among those proposing different cognitive architectures (Chi & Slotta, 1993; Chi, Slotta, & De Leeuw, 1994; diSessa, Gillespie & Esterly, 2004; Gupta, Hammer & Redish, 2010; Ioannides & Vosniadou, 2002; Slotta, Chi, & Joram, 1995). Understanding how concepts form in memory and how they are used to reason provides useful insights in devising instructional interventions to help students adopt and use scientific concepts.

Developing and evaluating instructional strategies to address students' misconceptions *What instructional interventions are most effective for helping students overcome stubborn misconceptions?* Since a primary goal of science instruction is to teach students the major concepts in each discipline as well as how to apply those concepts to solve problems, considerable effort has gone into research to design, evaluate, and refine curricular interventions to target students' stubborn misconceptions. This line of research builds upon the established catalogue of common misconceptions in the research literature and uses a cyclical process to devise instructional strategies. Most instructional techniques begin by making students aware of their misconceptions (e.g., by demonstrating to students inconsistencies in their own reasoning across contexts—see discussion of conceptual change below), and then guiding students through a series of activities or reasoning exercises to reshape their concepts to accommodate scientific concepts (Camp & Clement, 1994; Gupta et al., 2010; McDermott & Shaffer, 2002; Slotta &

Chi, 2006; Sokoloff & Thornton, 1997; Steinberg & Wainwright, 1993; Wosilait, Heron, Shaffer & McDermott, 1998).

Theoretical Framework

There are three main theoretical view points about conceptual development and conceptual change in the science education community, as summarized below.

Misconceptions View

This view contends that as students gain knowledge about the world (either through formal schooling or informally), they build “naïve theories” about how the physical world works, and that often these naïve theories contain misconceptions that contradict scientific concepts (Posner, Strike, Hewson, & Gerzog, 1982; Clement, 1982; McDermott, 1984; see Etkina, Mestre & O’Donnell, 2005 for a review). Hence, students do not come to class as “blank slates” upon which instructors can write appropriate scientific concepts (Bransford, Brown, & Cocking, 1999). For example, children observe leaves fluttering down from tree branches and rocks thrown from a bridge onto a stream below, and from these and multiple similar observations construct the notion that heavy objects fall faster than light objects. Misconceptions are viewed as stable entities that are used to reason about similar but varied contexts (e.g., a sheet of paper falling from a table to the floor falls slower than a set of keys falling to the floor, both observations fitting into “heavy objects fall faster than light objects” misconception). Misconceptions have three qualities: a) they interfere with scientific conceptions that teachers attempt to teach in science classes, b) they are deeply seated due to the time and effort that students spent constructing them, and they make sense to the student since they do explain many observations (it is difficult to “shut off” air resistance and observe a leaf and a rock falling at the same rate, thereby allowing direct verification of the physicist’s view that all objects fall at the same rate), and c) they are resistant to change.

Some researchers make distinctions between “weak” and “strong” restructuring of misconceptions. For example, young children’s eventual conception of the meaning of “alive” is considered strong restructuring (Carey, 1999); prior to correctly conceptualizing “alive” children believe that to be alive means being active (people and animals are alive by this criterion, but not lichen), or being real/seen (Carey conveys an incident whereby her 4-year-old daughter once proclaimed that it is funny that statues are dead even though you can see them but Grampa’s

dead because we can't see him any more—Carey, 1999, p. 306). Other distinctions include intentional versus non-intentional conceptual change, where the former is "...characterized by goal-directed and conscious initiation and regulation of cognitive, metacognitive, and motivational processes to bring about a change in knowledge" (Sinatra & Pintrich, 2003, p. 6).

In the misconceptions view it is generally accepted that some degree of dissatisfaction is needed for someone to replace a misconception with a more scientifically-appropriate form. As Carey argues, "Without doubt, the process of disequilibrium is an important part of the process of conceptual change" (Carey 1999, p. 306). Almost three decades ago Posner, Strike and their collaborators (Strike & Posner, 1982; Posner, et al., 1982) formulated a theory of conceptual change with four components that needed to be present for an individual to abandon a misconception in favor of a scientific concept: 1) Dissatisfaction with current concept, 2) Intelligibility of the new concept, 3) Initial plausibility of new concept, and 4) Usefulness of new concept in reasoning and making predictions about phenomena. Strike and Posner (1992) have since revised their initial views somewhat to include issues related to the learner's "conceptual ecology" as well as developmental and interactionist considerations.

Knowledge in Pieces or Resources View

Another view of students' knowledge and conceptual change offers a different architecture of concepts. According to the "knowledge in pieces" view (diSessa, 1993; diSessa & Sherin, 1998; Hammer, 1996), students' knowledge consists of smaller grain-size pieces that are not necessarily precompiled into larger concepts. Students activate one to several pieces of knowledge in response to context and reason with them on-the-fly. For example, individuals of all ages (Harvard-Smithsonian Center for Astrophysics, 1987) often state that it is hot during summer because the earth and the sun are in closer proximity than during winter. This is interpreted by misconception advocates as a misconception that likely formed from thinking about the path of the earth around the sun as an exaggerated ellipse. The pieces view is that perhaps the individual generated this explanation on the spot by searching memory for some reasonable explanation, and coming up with the knowledge piece "closer means stronger"; for example, moving closer to a fire makes the heat one feels radiating from it stronger, and moving closer to the stage in a rock concert makes the music louder (Hammer, 2000). The pieces view is similar to Hammer's "resources" view (Hammer, 2000), where resources are small units of

beliefs or thought whose “correctness” depends on the context in which they are applied, or to Minstrell’s (1992) “facets.”

Proponents of the pieces view argue that there are two major difficulties with the misconceptions view that are not present with the pieces view (diSessa, 1993; Smith, diSessa & Rochelle, 1994; Hammer, 1996a; 1996b; 2000; Hammer & Elby, 2002). The first is that a misconception is thought of as a compiled cognitive structure that an individual can employ to reason about a situation, which implies that the individual should show consistency in applying the misconception across similar contexts. Interviews with students as well as assessments reveal that slight changes in physics contexts can lead to radically different answers, reasoning patterns and/or explanations (Dufresne, Leonard, & Gerace, 2002; Steinberg & Sabella, 1997), suggesting that novice knowledge is highly context dependent and not stable as a misconceptions view would suggest. The second difficulty is that although misconceptions are constructed over time from experiences and observations, and therefore constructivist in nature, the misconceptions view does not provide an account of how misconceptions eventually evolve into correct scientific concepts. A recent publication, however, argues that the two camps are not as far apart as some might think (Brown & Hammer, 2008).

Ontological Categories View

A third perspective of novice knowledge and conceptual change is the “ontological categories” view. This view, attributed to Chi and her collaborators (Chi, 1992; 1997; Chi & Slotta, 1993; Chi, Slotta & de Leeuw, 1994; Slotta et al., 1995) and building on previous work by Kiel (1979), argues that students’ naïve conceptions is due to miscategorizing knowledge and experiences into inappropriate ontological categories, where ontological categories are loosely defined as the sorts of things there are in the world. Examples of ontological categories are material things like objects, temporal things like events, and processes like the emergent phenomenon that gives rise to flocking behavior. The ontological categories view argues that students sort knowledge into distinct and stable ontological categories and that many of the difficulties in student understanding are due to categorizing processes such as heat and electric current into the matter or “things” category. This would imply that the way to help students overcome misconceptions is to have them change their mis-categorized knowledge into the appropriate ontological categories, but misconceptions are difficult for students to dislodge precisely because it is hard for students to re-categorize across different ontological categories.

For example, students tend to think of forces as things rather than as the interaction between two (or more) objects, and hence talk of forces as being used up as gasoline is used up in a car. One well-known example of this is students discussing the forces on a ball thrown vertically up in the air—they mention “the force of the hand” as one of the forces “possessed” by the ball while it is rising, a force which gets used up when the ball reaches the top of its trajectory, after which the force of gravity takes over to make the ball fall (Clement, 1982).

Recently this view has been challenged (Gupta, Hammer & Redish, 2010). Gupta and his collaborators argue that both experts (in talks delivered to other experts in their profession) and novices (when discussing ideas with their peers) are able to traverse between ontological categories without confusion in order to explain complex phenomena. For example, Gupta et al. cite physics journal articles in which physicists mix ontological categories to build a coherent argument, such as discussing a pulse emerging from a sample and having a distinct peak as if it were a thing. In addition, they argue that novices do not show stable ontological categories, showing examples of novices speaking about electric current as a thing and one hour after instruction using much more sophisticated process-like descriptions of current. Finally, they also argue that crossing ontological categories is extremely common in how we communicate (e.g., phrases such as anger being “bottled up,” or having a cold that cannot “be shaken,” treating emotion and sickness as if they were things.)

Methodology (Data Collection/Sources and Data Analysis)

Contexts. Research on identifying misconceptions is done by either administering assessments designed to probe students views in contexts where concepts need to be applied or discussed, or by conducting clinical interviews with students about a context in which persistent conceptual errors seem to be prevalent. Clinical interviews has been the methodology most used in probing students conceptual architecture in memory; during clinical interviews, which typically last one hour, a researcher probes a student’s conceptual understanding in a target topic through a series of interviewer-led questions—questions that often are open-ended and guided by the student’s responses to previous questions. Because interview studies result in large amounts of data that are difficult to analyze and interpret, studies using this methodology rely on small numbers of students. Studies evaluating the efficacy of interventions for helping students overcome misconceptions are “engineered” (i.e., devised somewhat by trial and error based on best guesses

from research and teaching experience) using a cyclic process where initial best guesses are made in designing an initial intervention, then trying the intervention and evaluating its effectiveness, and then revising and starting the cycle again until learning outcomes are achieved.

Participants. Misconceptions research in STEM has been conducted with students of all ages and in various contexts such as large undergraduate introductory courses, as well as high school physics courses. Although not discussed in this synthesis, misconceptions on physical science concepts have also been studied with middle school and elementary school children. Recently there have been an increasing number of studies exploring conceptual understanding among upper-division physics students as well as graduate students.

Data Sources and Analysis Data sources for misconception studies come from students' performance in assessment questions or from transcripts of clinical interviews. In cases where an intervention or teaching approach is being evaluated, assessments of misconceptions are administered to students prior to, and following, the intervention and differences between the post and pre scores are analyzed. Data from interview studies are analyzed/interpreted using "grounded theory" (Strauss & Corbin, 1998) defined as developing a theory by observing the behavior of a group (in this case, students) in concert with the researcher's insights and experiences; any theory emerges from data as opposed to formulating hypotheses which are then tested.

Findings

Discussing research findings in conceptual change can be a controversial issue since, as discussed above, there are three theoretical perspectives describing the nature of concepts in memory and conceptual change. Misconceptions and ways of overcoming them are central themes according to the "misconceptions" theoretical view, whereas the knowledge in pieces or resources view is not in agreement with the robustness of misconceptions or with how one might go about "overcoming" them; proponents of this view would argue that as typically defined, students do not possess misconceptions, but rather compile knowledge pieces on-the-spot to reason about phenomena and thus the misconceptions that emerge for us to observe are highly

context dependent. There are also distinct differences between the ontological categories view and the other two in terms of explaining the nature of misconceptions and how to help students overcome them. It is therefore important to keep in mind these unresolved disagreements in the research community as some salient findings are presented below.

Misconceptions

- Students possess misconceptions that are deeply rooted and difficult to dislodge (Bransford, Brown, & Cocking, 1999; Etkina, Mestre, & O'Donnell, 2005). An abundance of misconceptions have been identified across a wide range of physics topics in undergraduate physics. Often misconceptions seemingly disappear and are replaced with scientific concepts following instruction only to reappear months later.
- Several intervention strategies, largely based on the Strike and Posner's (1982) conceptual change framework, have proven effective at helping students overcome misconceptions. Some strategies, such as one that has been successfully used by the University of Washington Physics Education Research group for many years, are based on a cyclic process that begins by indentifying misconceptions in a physics topic, then designing interventions based on previous research, instructor experiences or best guesses, then piloting the intervention and evaluating its success with post-tests that measure transfer to related situations, then refining the intervention over other cycles until evidence is obtained that the intervention works. It should be pointed out that even the most successful groups at this type of work (e.g., McDermott, Herron and Shafer at the University of Washington) will freely admit that devising effective instructional strategies to combat misconceptions is often a slow, painstaking task requiring multiple tries—not unlike engineering a solution to a complex problem. Other approaches, such as Clement's "bridging analogies," (Brown & Clement, 1989; Camp & Clement, 1994; Clement, 1993) start with "anchoring intuitions," which are strong and correct understandings that students possess, and devising interventions that attempt to bridge from students' correct intuitions to related contexts in which students display misconceptions. Yet another very effective approach (Sokoloff & Thornton, 1997) uses interactive lecture demonstrations to help students overcome prevalent misconceptions involving Newton's Third Law (i.e., the belief that heavy or fast moving objects exert larger forces on light or stationary objects when the two interact via collisions) by

displaying in real time the force exerted by mutually interacting carts during a collision under different conditions (e.g., heavy cart colliding with a light cart; moving cart colliding with a stationary cart, etc.). After experiencing the interactive lecture demonstrations intervention, students retain appropriate understanding of Newton's third law months afterwards.

- Despite the previous bullet, designing an effective intervention to help students overcome a particular misconception can be elusive (Harvard-Smithsonian Center for Astrophysics, 1987; 1997).

Knowledge in Pieces or Resources

- Students possess knowledge pieces, also referred to as “phenomenological primitives” (or p-prims), or resources, that are marshaled to reason about physics contexts. These pieces of knowledge can be recalled singly or in groups and compiled in different ways in real time in response to different contexts. It is common to observe that contexts that are considered equivalent, perhaps even nearly identical by experts are viewed as different by students and different knowledge pieces are brought to bear to reason about them. Students' knowledge is more dynamic in the pieces/resources view than in the misconceptions view. As expertise is developed with time and experience, there is refinement of the knowledge in memory whereby knowledge pieces that repeatedly prove effective to recall and apply to particular contexts are compiled into scientific concepts. That is, repeated rehearsal of knowledge pieces recalled and compiled to deal with similar situations leads to the formation of stable scientific concepts in experts.

Ontological Categories

- Misconceptions stem from students categorizing scientific ideas into inappropriate categories (Chi, 1992; 1997). For example, processes are placed in the things category (e.g., electric current is thought of as “fuel” that is used up in light bulbs).
- Misconceptions are relatively easy to fix when they involve modification within the same category, but are difficult to fix when they involve modifications across categories (Chi, 2005; Slotta, Chi & Joram, 1995).
- Instructional strategies designed to help students overcome misconceptions by recategorizing into appropriate ontological categories have shown promise (Chiu, Chou & Liu, 2002; Clark, 2006; Ferrari & Chi, 1998; Furio & Guisasola, 1998; Kikas, 2004;

Lee & Law, 2001; Slotta & Chi, 2006; Tuisson, Venville Harrison & Treagust, 1997; Venville & Treagust, 1998).

Strengths & Limitations of Conceptual Change Research

The strengths and limitations of this body of research include:

Strengths

- Misconceptions research has raised consciousness among instructors about students' learning difficulties, and about a misconception prevalent among instructors, namely that teaching done in a clear, elegant manner, even charismatic instructors quite often does not help students overcome misconceptions.
- Curricular interventions and assessments (both to be discussed in different sections) have emerged based on misconceptions research.
- Classroom instruction has changed as a result of misconceptions research, with many "active learning" strategies (e.g., the use of polling "clicker" technologies to teach large lecture courses) being practiced that have been shown more effective than traditional instruction at helping students overcome misconceptions.
- The pieces/resources and ontological categories views are attempting to map human cognitive architecture, which if successful can help in designing effective instructional strategies.

Limitations

- Designing "definitive" experiments that falsify one of the theoretical views remains elusive, hence the debate among proponents of the three theoretical views continues.
- Although many misconceptions have been catalogued across both introductory and advanced physics topics, it is daunting to ever achieve a complete list.

Areas for Future Study

Research on cataloguing misconceptions has been progressing for several decades and has covered an extensive range of physics topics (McDermott & Redish, 1999), so future research in this area is limited to alternate populations (e.g., upper-division students) and yet-to-be-investigated topics. There is a more pressing need for studies to help articulate general instructional strategies for guiding students to adopt scientific conceptions, especially when those

conflict with students' existing conceptions. Another promising area for future research is to design experiments to test the three competing viewpoints outlined in the theoretical framework in order to arrive at a more unified view.

References

- Ambrose, B.S., Heron, P.R.L., Vokos, S., & McDermott, L.C. (1999). Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena. *American Journal of Physics*, 67, 891-898.
- Beichner, R.J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62, 750-762.
- Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.
- Brown, D., & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: Factors influencing understanding in a teaching experiment. *Instructional Science*, 18, 237-261.
- Brown, D. E., & Hammer, D. (2008). Conceptual change in physics. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 127-154). New York: Routledge
- Camp, C. & Clement J. (1994). *Preconceptions in mechanics: Lessons dealing with students' conceptual difficulties*. Dubuque, Iowa: Kendall/Hunt.
- Carey, S. (1999). Sources of conceptual change. In E.K. Scholnick, K. Nelson, & P. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 293-326). Mahwah, NJ: Lawrence Erlbaum Assoc.
- Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science* (pp. 129-186). Minneapolis, MN: University of Minnesota Press.
- Chi, M. T. H. (1997). Creativity: Shifting across ontological categories flexibly. In T.B. War, S. M. Smith, & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery and change* (pp. 209-234). Washington, DC: American Psychological Association.
- Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, 14, 161-199.
- Chi, M.T.H. & Slotta, J.D. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction*, 10, 249-260.

- Chi, M. T. H., Slotta, J. D., de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27-43.
- Chiu, M., Chou, C., & Liu, C. (2002). Dynamic processes of conceptual changes: Analysis of constructing mental models of chemical equilibrium. *Journal of Research Science Teaching, 39*, 688-712.
- Clark, D.B. (2006). Longitudinal conceptual change I students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction, 24*, 467-563.
- Clement, J.J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*, 66-71.
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30*(10), 1241-1257.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition & Instruction, 10*(2-3), 105-225
- diSessa, A. A., Gillespie, N., & Esterly, J. (2004). Coherence vs. fragmentation in the development of the concept of force. *Cognitive Science, 28* (6), 843-900.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*(10), 1155-1191
- Dufresne, R.J., Leonard, W.J., & Gerace, W.J. (2002). Making sense of students' answers to multiple choice questions. *The Physics Teacher, 40*, 174-180.
- Etkina, E., Mestre, J. & O'Donnell, A. (2005). The impact of the cognitive revolution on science learning and teaching. In J.M Royer (Ed.) *The Cognitive revolution in educational psychology* (pp. 119-164). Greenwich, CT: Information Age Publishing.
- Ferrari, M., & Chi, M. T. H. (1998). The nature of naïve explanations of natural selection. *International Journal of Science Education, 20*, 1231-1256.
- Furio, C., & Guisasola, J. (1998). Difficulties in learning the concept of electric field. *Science Education, 82*, 511-526.
- Goldberg, F.M. & McDermott, L.C. (1986). Student difficulties in understanding image formation by a plane mirror. *The Physics Teacher, November*, 472-480.
- Goldberg, F.M. & McDermott, L.C. (1987). An investigation of student understanding of the real image formed by a converging lens or concave mirror. *American Journal of Physics, 55*, 108-119.

Gupta, A., Hammer, D. & Redish, E.F. (2010). The case for dynamic models of learners' ontologies in physics. *Journal of the Learning Sciences*, 19(3), 285-321.

Hammer, D. (1996a). Misconceptions or p-prims: How might alternative perspectives on cognitive structure influence instructional perceptions and intentions. *Journal of the Learning Sciences*, 5, 97-127.

Hammer, D. (1996b). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, 64, 1316-1325.

Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics Physics Education Research Supplement*, 68(7) S52-S59.

Hammer, D & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer and P. R. Pintrich, (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp 169-190). Mahwah, NJ: Erlbaum.

Harvard-Smithsonian Center for Astrophysics (Producer). (1987). *A private universe* [Video documentary]. Retrieved September 15, 2010, from <http://www.learner.org/resources/series28.html>

Harvard-Smithsonian Center for Astrophysics (Producer). (1997). *Minds of our own: Lessons from thin air* [Video documentary]. Retrieved September 15, 2010, from <http://www.learner.org/resources/series26.html>

Helm, H. & Novak, J.D., (Eds.). (1983). *Proceedings of the International Seminar on Misconceptions in Science and Mathematics*. Ithaca, NY: Department of Education, Cornell University.

Hestenes, D. & Wells, M. (1992). A mechanics baseline test. *The Physics Teacher*, 30, 141-158.

Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30, 159-166.

Ioannides, C., & Vosniadou, S. (2002). The changing meanings of force. *Cognitive Science Quarterly*, 2(1), 5-62.

Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Kikas, E. (2004). Teachers' conceptions and misconceptions concerning three natural phenomena. *Journal of Research in Science Teaching*, 41, 432-448.

Lee Y., & Law, N. (2001). Explorations in promoting conceptual change in electrical concepts via ontological category shift. *International Journal of Science Education*, 23, 111-149.

Maloney, D.P., O'Kuma, T.L., Hieggelke, C.J., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. *American Journal of Physics, Physics Education Research Supplement*, 69(7), S12-S23

McDermott, L.C. (1984). Research on conceptual understanding in mechanics. *Physics Today*, 37(7), 24-32.

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

McDermott, L.C. & Shaffer, P.S. (1992a). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics*, 60, 994-1003.

McDermott, L.C. & Shaffer, P.S. (1992b). *Tutorials in Introductory Physics*. Prentice Hall: Upper Saddle River, NJ.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In R. Duit, F. Goldberg & H. Niedderer (Eds.), *The Proceedings of the International Workshop on Research in Physics Education: Theoretical Issues and Empirical Studies* (pp. 110-128) (Bremen, Germany, March 5-8, 1991). Kiel, Germany: IPN (Institut für die Pädagogik der Naturwissenschaften).

Novak, J.D. (Ed.). (1987). *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Vols. I, II & III*. Ithaca, NY: Department of Education, Cornell University.

O'Kuma, T.L., Maloney, S.P., & Hieggelke, C.J. (2000). *Ranking task exercises in physics*. Upper Saddle River, NJ: Prentice Hall.

Posner, G., Strike, K., Hewson, P. & Gerzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211-227.

Sinatra, G. M. & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. In G. M. Sinatra & P. R. Pintrich (Eds.). *Intentional Conceptual Change* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum Associates.

Slotta, J., & Chi, M. T. H. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, 24, 261-289.

Slotta, J. D., Chi, M. T. H. & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, 13, 373-400.

- Smith, J., diSessa, A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, **3** (2), 115-163.
- Sokoloff, D.R., & Thornton, R.K. (1997). Using interactive lecture demonstrations to create an active learning environment. *The Physics Teacher*, **35**(6), 340-347.
- Steinberg, R.N. & Sabella, M.S. (1997). Performance on multiple-choice diagnostics and complementary exam problems. *The Physics Teacher*, **35**(March), 150-155.
- Steinberg, M.S. & Wainwright, C.L. (1993). Using models to teach electricity--The CASTLE project. *The Physics Teacher*, **31**(6), 353-357.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research*. Thousand Oaks, Calif.: Sage.
- Streveler, R. A., Litzinger, T. A., Miller, R. L., & Steif, P. S. (2008). Learning conceptual knowledge in the engineering sciences: overview and future research directions *Journal of Engineering Education*, **97**(3), 279–294.
- Strike, K.A., & Posner, G.J. (1982). Conceptual change and science teaching. *International Journal of Science Education*, **4**(3), 231-240.
- Strike, K.A. & Posner, G.J. (1992). A revisionist theory of conceptual change. In R.A. Duschl & R.J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147-176). New York: State University of New York Press.
- Thornton, R.K. & Sokoloff, D.R. (1998). Assessing student learning of Newton's Laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula, *American Journal of Physics*, **66**, 338-352.
- Tyson, L. M., Venville, G. J., Harrison, A. G., & Treagust, D. F. (1997). A multidimensional framework for interpreting conceptual change events in the classroom. *Science Education*, **81**, 387-404.
- Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching*, **35**, 1031-1055.
- Wosilait, K., Heron, P.R.L., Shaffer, P.S. & McDermott, L.C. (1998). Development and assessment of a research-based tutorial on light and shadow. *American Journal of Physics*, **66**, 906-913.

Problem Solving

In addition to research on conceptual understanding, another key focus of physics education research on student learning is *problem solving*, likely because problem solving is a key component of most physics courses. As with other studies of student learning, this research focuses first on documenting what students do while solving problems, and follows with the development and evaluation of instructional strategies to address student difficulties. Since problem solving is a complex cognitive process, this area of research also has a strong overlap with the section on *Cognition*. For additional reviews of research on physics problem solving, see Maloney (1994) and Hsu, Brewster, Foster, & Harper (2004).

Research Questions

Key areas of research on physics problem solving include expert-novice research, worked examples, representations, mathematics in physics, and evaluating the effectiveness of instructional strategies for teaching problem solving.

Expert-novice research. *What approaches do students use to solve physics problems? How are the problem solving procedures used by inexperienced problem solvers similar to and different from those used by experienced solvers? How do experts and novices judge whether problems would be solved similarly?* Early studies of physics problem solving investigated how beginning students solve physics problems and how their approaches compare to experienced solvers, such as professors (Bagno & Eylon, 1997; Chi, Glaser, & Rees 1982; Eylon & Reif, 1984; Larkin 1979, 1981a, 1981b; Larkin, McDermott, Simon, & Simon, 1980). This area of research also includes categorization studies to infer how physics knowledge is structured in memory (Chi, Feltovich, & Glaser, 1981; Hardiman, Dufresne, & Mestre, 1989).

Worked examples *How do students study worked-out examples? How do students use solutions from previously solved problems when solving new problems? How do students use instructor solutions to find mistakes in their problem solutions to homework and exams? What features of worked-out problem solutions facilitate student understanding of the example?*

This body of research explores how students use worked-out problem solutions or previously solved problems to solve new unfamiliar problem(s) (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Ferguson-Hessler & de Jong, 1990). It also includes how students use instructor solutions to self-diagnose errors in their own problem solutions (Cohen, Mason, Singh, & Yerushalmi, 2008; Yerushalmi, Mason, Cohen, & Singh, 2008), and how to design effective examples (Ward & Sweller, 1990).

Representations. *What representations are constructed and used by students during problem solving? How are representations used by students? What is the relationship between facility with representations and problem solving performance? What instructional strategies promote students' use of representations?* This research explores the use of external representations for describing information during problem solving, such as pictures, physics-specific descriptions (e.g. free-body diagrams, field line diagrams, or energy bar charts), concept maps, graphs, and equations. Some studies focus on what representations are constructed during problem solving and the manner in which they are used (Heller & Reif, 1984; Larkin, 1983; Rosengrant, Van Heuvelen, & Etkina, 2005; Van Heuvelen & Zou, 2001), whereas other studies explore the facility with which students or experts can translate across multiple representations (Kohl & Finkelstein, 2005; Meltzer, 2005).

Mathematics in physics *How does the application of mathematics in a physics course differ from the mathematical skills taught in a math course? How do students interpret and use symbols during quantitative problem solving?* This area of research explores how quantitative tools from mathematics courses are applied during physics problem solving. Some examples include the use of symbols in equations to represent physical quantities (Sherin, 1996; 2001; Torigoe, 2008; Tuminaro & Redish, 2007), vector addition (Nguyen & Meltzer, 2003), arithmetic, algebra, geometry, calculus (e.g. integration) (Cui, Rebello, & Bennett, 2006), and proportional reasoning (Cohen & Kanim, 2005).

Evaluating the effectiveness of instructional strategies for teaching problem solving *How does instructional strategy X [e.g., cooperative group problem solving] affect students' problem solving skills? To what extent do conceptual approaches to problem solving influence students'*

conceptual understanding of physics? How do students interact with online computer tutor systems? What features of web-based homework systems successfully enhance students' problem solving skills? Several instructional strategies have been developed and tested, including the use of alternate types of problems (Heller & Hollabaugh, 1992; Mestre, 2002; O'Kuma, Maloney, & Hieggelke, 2000; Van Heuvelen, 1995; Van Heuvelen & Maloney, 1999), adopting an explicit problem solving framework (i.e. a consistent sequence of problem solving steps) (Wright & Williams, 1986), conceptual approaches such as writing qualitative strategies (Leonard, Dufresne, & Mestre, 1996; Van Heuvelen, 1991b), cooperative group problem solving (Heller, Keith, & Anderson, 1992), and computer homework or tutor systems to help students become better problem solvers (Kortmeyer et al., 2008; Lee, Palazzo, Warnakulassoriya, & Pritchard, 2008; Reif & Scott, 1999).

Theoretical Frameworks

This section identifies a few prominent theories about learning and problem solving from cognitive science and educational psychology. Although these frameworks are viewed as useful and relevant for PER, problem solving researchers often do not clearly define, or draw upon a theoretical basis for their research studies. The frameworks reviewed here are intended to provide a starting point for discussion. The following frameworks are included: Information-Processing Models, Problem Solving by Analogy, Resources Model, and Situated Cognition.

Information-Processing Models

According to one theory of human problem solving (Newell, Shaw, & Simon, 1958; Newell & Simon, 1972; Simon, 1978), problem solving is an iterative process of representation and search for a solution. They define a mental state called the "problem space" consisting of a person's available knowledge and their internal representation or understanding of the task environment, including the initial state (given information), goal state or target, and appropriate operations that can be employed. After representing the problem, a solver engages in a search process during which s/he selects a goal and a method to apply (such as a general heuristic), applies it, evaluates the result of this choice, modifies the goal or select a subgoal, and proceeds

in this fashion until a satisfactory solution is achieved or the problem is abandoned. Gick (1986) expanded upon this model to include the role of schema activation. If constructing an internal representation of the problem activates an existing schema or memory framework for solving the problem, a solution is implemented immediately. If no schema is activated, then the solver engages in general search strategies (Gick, 1986).

Within this theory of problem solving, Newell and Simon (1972) also describe the organizational structure of memory as an information-processing system (IPS) consisting of two branches: short-term or “working” memory (STM) and long-term memory (LTM). Short-term memory is constrained and can only hold a small amount of information for a limited time, whereas long-term memory is essentially unlimited (Simon, 1978). However, in order to access information stored in long-term memory it must be activated and brought into working memory. If problem information and activated knowledge exceeds the limits of STM, a solver may experience *cognitive overload* (Sweller, 1988) which interferes with attempts to reach a solution. To alleviate this effect, problem information is often stored externally (e.g., written down on paper) or processed with the aid of a tool (e.g., computer) in order to free up space in working memory to devote to the task. As a problem-solver becomes more proficient, knowledge and procedures may become *chunked* and some skills become automatic, which also optimizes the capacity of STM.

Problem Solving by Analogy

Instructors report that students often look for similar example problems in physics textbooks when attempting to solve homework problems. This process of using a similar, familiar problem to solve a new problem is referred to as *analogy*. There are three main views of analogical transfer, including structure-mapping (Gentner, 1983; 1989), pragmatic schema view (Holyoak 1985; Holyoak & Koh, 1987), and exemplar view (Ross 1987, 1989). These views agree that there are three main criteria for making use of a known problem: 1. a person must be reminded of the previous problem and sufficiently remember its content, 2. s/he must compare and adapt the attributes from one problem to the other (what Gentner calls structure-mapping), and 3. the solver must evaluate the effectiveness of this transfer for solving the new problem (Reeves & Weisberg, 1994). To be useful for future problem solving, a person must also abstract common features and generalize from the examples. The work of Holyoak and Koh (1987) suggests that similarity can be both at a surface level and at a structural level, and while surface

similarity can aid with reminding, it is structural similarity that is important for appropriately solving the new problem. Exemplar-based models emphasize the role of specific example problems for guiding the application of abstract principles (Ross, 1987; 1989). For a more comprehensive review of these views see Reeves & Weisberg, 1994.

Resources Model & Epistemic Games

Research on how students use and understand mathematical symbols and equations in physics has built upon a theoretical framework called the resources model (Hammer, 2000; Hammer, Elby, Scherr, & Redish, 2005). This model describes knowledge in terms of students' resources for learning, both "conceptual" and "epistemological" in nature (Hammer, 2000). This model helps to explain why students might have the requisite knowledge and skills for solving a problem but fail to activate it in a particular context, and suggests that instruction should be designed to make productive use of students' resources. Tuminaro (2004) and Tuminaro and Redish (2007) extended this resources model to give a detailed framework for analyzing students' application of mathematics to physics problems. They identified six *epistemic games* or problem solving approaches that students participate in while using math in physics, including: mapping meaning to mathematics, mapping mathematics to meaning, physical mechanism, pictorial analysis, recursive plug-and-chug, and transliteration to mathematics.

Situated Cognition

Some research on instructional strategies for teaching problem solving (e.g. cooperative problem solving) is based upon the situated cognition model of learning. In this view, knowledge is "situated" or influenced by the particular task and context in which learning is taking place (Brown, Collins, & Duguid, 1989). The authors propose that school activities should incorporate authentic tasks, and they propose a teaching method called cognitive apprenticeship to facilitate an enculturation process. Cognitive apprenticeship includes three main phases: modeling (demonstrating how to do something), coaching (providing opportunities for guided practice and feedback), and fading (gradually removing scaffolding).

Methodology (Data Collection/Sources and Data Analysis)

Contexts. Basic research to understand how students solve problems and identify common student difficulties is typically done outside the classroom, in an experimental or clinical setting

with paid volunteers. Studies of pedagogy and curricular interventions are conducted in conjunction with a course where the data collected includes student responses to course materials such as homework and exams.

Participants. Studies of physics problem solving frequently involve undergraduate students in large-scale introductory courses (algebra-based or calculus-based) at a single institution over one or two semesters, sometimes longer. Studies conducted with high school students or disciplinary majors are less common. The “experts” in expert-novice studies are often physics faculty or graduate teaching assistants who are experienced with teaching introductory courses.

Data Sources and Analysis One common data source for problem solving research is students’ written solutions to physics problems which have been designed by the researcher(s) and/or adapted from existing problem sources. In most cases the problems are free response but occasionally they are in a multiple-choice format. This data is typically analyzed by scoring the solutions according to particular criteria, as determined by comparing the written solutions to features of an “ideal” instructor solution. Sometimes rubrics are used to make scoring criteria objective and improve agreement among multiple scorers. Students’ written solutions are used in curricular intervention studies to compare performance in reformed courses to nonreformed courses on a set of common assessment items.

Another widespread data source, especially for cognitive studies, is think-aloud problem solving interviews. In these studies participants are asked to solve physics problems and verbalize their thoughts during the process while being video and/or audio taped. Transcripts from these interviews are analyzed using standard qualitative methods such as case studies or grounded theory to elucidate common themes or problem solving approaches from the statements.

Studies comparing expert and novice problem solvers have traditionally used categorization tasks that require participants to group problems based on similarity of solution. Early studies used card-sorting tasks where subjects physically placed problem statement cards into piles and the resulting category was assigned a name (Chi et al., 1981; Mason & Singh, 2009). More recent studies have used alternate categorization tasks, such as selecting which of two problems would be solved most like a model problem (Hardiman, Dufresne & Mestre, 1989)

multiple choice categorization formats (Leonard, Dufresne, & Mestre 1996) or ranking the similarity of two problems on a numerical scale (Mateycik, Rebello, & Jonassen, 2010). Some categorization tasks are analyzed using qualitative methods (to elicit commonalities across card groupings) whereas others compare quantitative ranking or the frequency of similarity judgment responses for different groups of subjects.

Eye-tracking technology is a rare data source, but in some cases it is used to study gaze patterns or fixation time upon particular features of problems or problem solutions presented on a screen (Carmichael, Larson, Gire, Loschky, & Rebello, 2010; Smith, Mestre & Ross, 2010). These problem features could include representations (such as pictures or physics diagrams), text, equations, and mathematical steps in example solutions.

Findings

Expert-Novice Studies

Early studies of problem solving identified differences between beginning problem solvers and experienced problem solvers both in the way they organize their knowledge of physics and how they approach problems.

- *Categorization.* Expert-novice studies found that when asked to group problems based on similarity of their solution, beginning problem solvers used literal objects from the surface attributes of the problems as category criteria (such as “spring” problems or “incline plane” problems) whereas experienced problem solvers considered the physics concept or principle used to solve the problem when deciding on problem categories, such as grouping Conservation of Energy problems together (Chi, Feltovich, & Glaser, 1981; de Jong & Ferguson-Hessler, 1986). It was later shown that expert-novice is a continuum not a dichotomy; sometimes beginning students exhibit expert-like behaviors and experienced solvers behave like novices (Hardiman, Dufresne, & Mestre, 1989). Further, this same study also demonstrated that novices also rely on terminology and variable names in the problems to make categorization decisions, not just objects in the problem.
- *Problem-solving approaches.* Think-aloud problem solving interviews found that expert problem-solvers typically begin by describing problem information qualitatively and using

that information to decide on a solution strategy before writing down equations (Bagno & Eylon, 1997; Chi, Glaser, & Rees 1982; Eylon & Reif, 1984; Larkin 1979, 1981a, 1981b; Larkin, McDermott, Simon, & Simon, 1980). A successful solver's strategy usually includes the appropriate physics concept or principle and a plan for applying the principle to the particular conditions in the stated problem (Finegold & Mass, 1985). In contrast, beginning physics students typically started by writing down equations that match given or desired quantities in the problem statement and performing mathematical manipulations to get an answer (Chi et al., 1981; Walsh, Howard, & Bowe, 2007).

- *Metacognition*. In addition, these studies concluded that experts monitor their progress while solving problems and evaluate the reasonableness of the answer, whereas beginning students frequently get stuck and lack strategies to progress further (Chi, 2006; Chi et al. 1989; Larkin, 1981; Reif & Heller, 1982; Singh, 2002). Explicit instruction on strategies for evaluating a solution, such as the use of limiting and extreme cases, has been shown to improve performance (Reif & Heller, 1982).

Worked Examples

A common practice in problem solving instruction is to provide students with example problems that demonstrate solution steps. This section briefly reviews research on how students study example problem solutions in textbooks, how they refer to previous examples during problem solving, and how they use instructor solutions to detect and correct errors in their own solutions. This research also addresses how to design example solutions to optimize their usefulness.

- *Using examples*. Research conducted by Chi et al. (1989) and Ferguson-Hessler & de Jong (1990) on how students study worked-out example solutions in textbooks concluded that good and poor students used worked examples differently when solving a new problem. In Chi et al. (1989), the labels “good” and “poor” were determined by a student's success at solving problems after studying the worked examples (12 isomorphic / similar problems and 7 unrelated problems from a textbook). Good students (who scored an average of 82% on problems) referred to a specific line or line(s) in the example to check procedural aspects of their solution, whereas poor students (who only scored an average of 46% on problems) reread the example from the beginning to search for declarative information and a solution procedure they could copy (Chi et al., 1989). Ferguson-Hessler & de Jong (1990) confirmed

these findings and described the actions of good students as “deep processing” of the examples whereas poor students engaged in “superficial processing.”

A study by Smith, Mestre, & Ross (2010) used eye tracking to investigate what aspects of a problem solution students look at while studying worked-out examples. They found that although students spent a large fraction of time reading conceptual, textual information in the solution, their ability to recall this information later was poor. The students’ eye-gaze patterns also indicated they frequently jumped between the text and mathematics in an attempt to integrate these two sources of information.

- *Self-diagnosis*. When students receive feedback on their performance on homework and exams, instructors generally expect students to reflect on and learn from their mistakes. However, research indicates that such self-diagnosis of errors is difficult for students. In a study by Cohen, Mason, Singh, and Yerushalmi (2008) only one-third to one-half of students were able to recognize when they used an inappropriate physics principle for a problem. These results were affected by the support provided to the students, indicating that having access to a correct solution and self-diagnosis rubric helped students explain the nature of their mistakes better than an instructor solution alone or only access to the textbook and class notes (Yerushalmi, Mason, Cohen, & Singh, 2008). A classroom intervention by Henderson and Harper (2009) found that requiring students to correct mistakes on assessments improved student learning of content and also improved skills relevant for problem solving, such as reflection.
- *Example structure and the worked-example effect*. Ward and Sweller (1990) conducted classroom experiments using topics of geometric optics and kinematics to investigate how the design of example solutions influences learning and problem solving. They found that under certain conditions, reviewing correct example solutions was more effective for learning than actually solving problems, what is referred to as the *worked-example effect*. To optimize their usefulness, they claim that examples must direct students’ attention to important problem features and minimize cognitive load by keeping information sources separated, such as diagrams, text, and equations.

Representations

The term *representation* has multiple interpretations, but for the problem solving research reviewed here it is used to refer only to concrete, external descriptions used by a solver. Some examples include pictures or sketches, physics-specific descriptions (e.g. free-body diagrams, field line diagrams, ray diagrams, or energy bar charts), concept maps, graphs, and equations / symbolic notation. Some researchers go on to make a distinction between general and physics-specific representations. Reif and Heller (1982) suggest that *abasic* description includes initial steps taken to understand a problem, such as introducing symbolic notation and summarizing problem information verbally or pictorially. They separate this from *theoretical* description that specifies systems and interactions for objects using physics concepts, such as describing motion with position, velocity, and acceleration or describing interactions by force vectors.

- *Representational format affects performance.* Multiple studies have determined that some students give inconsistent answers to the same problem solving question when it is presented in different representational formats (Kohl & Finkelstein, 2005; Meltzer, 2005). The four formats tested included verbal, diagrammatic (or pictorial), mathematical/symbolic, and graphical. The pattern of performance differed by the task and problem topic, but performance on mathematical quizzes was usually worse than other formats despite students' preference for calculation questions (Kohl and Finkelstein, 2005).
- *Physics-specific descriptions (might) facilitate problem solving.* A clinical study by Heller and Reif (1984) found that subjects who received training materials on procedures for generating physics-specific, theoretical descriptions in mechanics performed better on a set of three problems than an unguided group. The group scored significantly higher on all four aspects of their solutions: describing the motion of objects, drawing a force diagram, correct equations, and the final answer. The authors concluded that these “descriptions markedly facilitate the subsequent construction of correct problem solutions” (p. 206). In a study by Larkin (1983), five students who were trained to use physical representations for direct-current circuits solved significantly more problems on average than a group of five students who received training on generating and combining equations to obtain an answer. The subjects were presented with three problems, and a problem was considered “solved” if the student produced a correct solution within a 20-minute time limit. Other studies have produced mixed results regarding the relationship between representations and problem

solving success. Studies by Rosengrant, Van Heuvelen, and Etkina (2005; 2009) found that students who drew correct free-body diagrams were much more likely to solve a problem correctly than those who drew an incorrect diagram or none at all, but those who drew an incorrect diagram performed worse than students who drew no diagram. In a study by Van Heuvelen and Zou (2001), students reported multiple work-energy representations were useful tools for solving problems, but the researchers observed little use of the representations on exams and could not determine their relationship to problem solving performance. Similarly, Rosengrant and Mzoughi (2008) found favorable student reactions to the use of impulse-momentum diagrams during class, but no differences in performance. De Leone and Gire (2006) did not find a significant correlation between the number of non-mathematical representations used by a problem solver and the number of correct answers, but the nature of the errors made was different for representation users and non-users.

- *Successful problem solvers use representations differently than unsuccessful solvers* Some studies indicate that successful problem-solvers use qualitative representations to guide the construction of quantitative equations (Ploetzner et al., 1999) and they use physics diagrams as a tool to check the consistency of their work (Rosengrant, Van Heuvelen, & Etkina, 2006). In contrast, unsuccessful solvers either do not generate representations or do not make productive use of them for problem solving. Kohl and Finkelstein (2008) suggest that both novices and expert generate multiple representations during problem solving and sometimes do so in a different order, but experts can more flexibly move between those multiple representations.

Mathematics in Physics

Mathematics and equations are often referred to as the “language of physics”. This area of study identifies common student difficulties associated with math skills required for physics problems and describes how equations are used differently in the context of physics than in math.

- *Vectors* Studies have documented several student difficulties associated with using vectors, particularly for velocity, acceleration, forces, and electric fields (Aguirre & Erickson, 1984; Flores, Kanim, & Kautz, 2004; Nguyen & Meltzer, 2003; Shaffer & McDermott, 2005). In particular, Nguyen and Meltzer (2003) identified widespread difficulties with two-

dimensional vector addition and interpreting the magnitude and direction of vectors in graphical problems.

- Algebraic or proportional reasoning. Some students have difficulty translating a sentence into a mathematical expression, and in doing so they often place quantities on the wrong side of an equals sign. Cohen and Kanim (2005) studied occurrences of the algebra “reversal error” in both algebra and physics contexts, concluding that sentence structure was the most influential factor in guiding translation of sentences expressing relationships among variables into equations.
- Symbols and equations. In contrast to mathematical expressions, equations used in physics have conceptual meaning associated with symbols and the relationships among physical quantities (Tuminaro, 2004; Tuminaro & Redish, 2007). Students generally have poorer performance on symbolic questions than on numeric questions, and Torigoe (2008) attributes this difference to students’ confusion about the meaning of symbols in equations and difficulty keeping track of multiple quantities. Sherin (1996, 2001) identified 21 different symbolic forms for equations in physics and how students interpret them, concluding that it is possible for students to have mathematically correct solutions but inappropriate conceptual understanding for the equations.
- Use of mathematical skills is context-dependent. In many cases, students have the mathematical resources to solve problems but do not attempt to apply them in the context of a physics class or vice versa (Bassok and Holyoak, 1989; Meredith & Marrongelle, 2008; Sabella & Redish, 2007). This suggests that the availability of resources depends on the context in which a skill was learned and the solver’s perception of appropriate knowledge to activate in a particular situation, referred to as *epistemological framing* (Bing & Redish, 2009; Tuminaro, 2004). Cui, Rebello, and Bennett (2006) found that students needed prompting and guidance to make the connection between their knowledge of calculus and a physics problem. Sayre & Wittman (2008) found that students had a more solid understanding of Cartesian coordinate systems and would persist in applying this system even when a different (polar) coordinate system was more appropriate. They used *Resource Theory* and *Process/Object Theory* to explain the development of unfamiliar ideas, suggesting a “plasticity continuum” measure in which resources can range from being

plastic (uncertain about how and when to apply them) to being solid (familiar and confident in their understanding of how to apply them).

Instructional Strategies

- *Alternate problem types.* Standard problems presented in textbooks are often well-defined and arguably do not reflect the nature of scientific thinking (Maloney, 1994). Several types of problems have been developed as alternatives to these standard problems which emphasize conceptual reasoning for realistic contexts. Some examples include context-rich problems (Heller & Hollabaugh, 1992; Ogilvie, 2009), experiment problems (V an Heuvelen, 1995; V an Heuvelen, Allen, & Mihas, 1999), jeopardy problems (V an Heuvelen & Maloney, 1999), problem posing (Mestre, 2002), ranking tasks (Maloney, 1987; O’Kuma, Maloney, & Hieggelke, 2000), real-world problems (<http://relate.mit.edu/RwProblems/>), and thinking problems (<http://www.physics.umd.edu/perg/abp/think/>).

The effects of implementing alternative problems are sometimes difficult to separate from the instructional strategies used in conjunction with the problems. For example, context rich problems (Heller & Hollabaugh, 1992; Heller, Keith, & Anderson, 1992) result in improved problem solving performance when used in conjunction with a problem solving strategy and cooperative group problem solving. Specifically, students made fewer conceptual mistakes, generated more useful descriptions, and wrote equations that were consistent with their descriptions.

Experiment problems are typically utilized during laboratory sessions, where students can explore the question using objects similar to those in the problem (V an Heuvelen, 1995). The article only describes the problems, and does not provide research results on their effectiveness. The use of Jeopardy Problems resulted in high scores on the Mechanics Baseline Test and the force Concept Inventory (V an Heuvelen & Maloney, 1999), but the authors caution that other instructional curricula were also used in conjunction with these problems.

In problem posing (Mestre, 2002), students must generate a problem statement that meets certain criteria (for example, they are shown a diagram of two blocks connected by a string and they must pose a problem that can be solved by Newton’s second law). This was found to be an effective tool for assessing a student’s knowledge of physics concepts and their

problem-solving skills, especially their ability to link appropriate problem contexts with physics principles.

- *Problem solving frameworks.* Explicitly modeling an organized set of problem solving steps and reinforcing this framework throughout a course results in higher course performance (Huffman, 1997; Heller & Reif, 1984; Wright & Williams, 1986; Van Weeren et al., 1982). Often such frameworks are based on the work of the mathematician Polya (1945): understand the problem, plan the solution, execute the plan, and look back (Heller & Heller, 2000; Reif, 1995; Van Heuvelen, 1991a) but are expanded to include explicit guidelines for constructing an initial description of the problem.
- *Qualitative approaches.* Some curricula emphasize performing a qualitative analysis of a problem before writing down quantitative equations. Some examples include OCS (Overview, Case Study Physics) (Gautreau & Novemsky, 1997; Van Heuvelen, 1991a; 1991b), hierarchical problem analyses based on principles (Dufresne, Gerace, Hardiman, & Mestre, 1992; Mestre et al., 1993) and strategy-writing where a strategy consists of the principle, justification, and procedure needed for solving a problem (Leonard, Dufresne, & Mestre, 1996).
- *Problem solving instruction.* Some studies describe how courses can be restructured to include explicit instruction on problem solving skills, such as through the use of cooperative group problem solving (Heller, Keith, & Anderson, 1992) or a curriculum called ALPS: Active Learning Problem Sheets (Van Heuvelen 1994a, 1994b). For additional instructional strategies, see a meta-analysis presented in Taconis, Ferguson-Hessler, and Broekkamp (2001).
- *Computer tutors.* Early uses of computers as problem solving tutors focused on prompting students to use a systematic approach for solving Newton's laws and kinematics problems (Kane & Sherwood, 1980; Sherwood, 1971; Smith & Sherwood, 1976). These tutoring systems have expanded substantially to include sophisticated hints, guidance, and feedback (Gertner & VanLehn, 2000; Hsu & Heller, 2009; Reif & Scott, 1999; Scott, 2001). Other computer-assisted programs focus on web-based homework such as the "computer-assisted personalized assignment" system CAPA (Kashy et al., 1993; Kashy et al. 1995; Kortmeyer et al., 2008; Pascarella, 2002), Mastering Physics, and CyberTutor (Lee, Palazzo, Warnakulassoriya, & Pritchard, 2008).

Studies comparing paper-and-pencil homework to web-based homework have found mixed results, with some studies citing no differences in performance (Bonham, Deardorff, & Beichner, 2003) and others citing improved performance with web-based tutors or homework systems (Cheng, K.K., Thacker, B.A., Cardenas, R.L., & Crouch, C., 2004; Dufrene, Mestre, Hart & Rath, 2002; Morote & Pritchard, 2009; VanLehn et al., 2005). Additional research results about homework are summarized in the section on *Assessment*.

Strengths & Limitations of Problem Solving Research

Strengths. A strength of problem solving research is that it is an established subfield of PER with a strong history and associations to other fields, such as mathematics and cognitive science. As a result, we have a great deal of information about how students solve problems and instructional strategies to address common difficulties. In addition, we can make use of established research methodologies from cognitive science and psychology when conducting research in physics education.

Limitations. Many of the conclusions from problem solving research have been drawn from studies with low number of subjects, a limited range of problem types and topics, and inconsistent measures of problem solving performance. In addition, the complex nature of problem solving and variability in problem features make it difficult to isolate particular factors that may be responsible for improvements (or decrements) in problem solving performance.

- Low numbers of subjects. Some early studies had very few subjects and/or the results have not been replicable on a larger scale (see for example, Mason & Singh's 2009 paper on revisiting categorization). Also, few studies have investigated problem solving approaches of transition stages in the expert-novice continuum, such as undergraduate majors and graduate students (Gire, 2007; Harper, 2001).
- Limited range of problem solving tasks. In many expert-novice studies, the experts are solving exercises that are simple for them, not novel problems they have not seen before. Very few studies have investigated the approaches that experts take on complex problems (see for example Singh, 2002). There is little research comparing problem solving across

physics topics, for example contrasting problem solving in mechanics with problem solving in electricity and magnetism or solving problems in upper-division physics courses.

- *Inconsistent problem solving measures.* There is no widespread, consistent measure of problem solving performance (Henderson et al., 2004). In most of the studies reviewed here, problem solving was measured using vague techniques such as mean number of correct answers, exam scores, or course grades without presenting detailed information about the way in which problem solutions were scored. There is some current research to develop a valid and reliable measure of problem solving, but for now this is still a limitation of many studies (Adams & Wieman, 2006; Marx & Cummings, 2010; Docktor, 2009).
- *Failure to systematically consider problem features.* Problem solving is a complex topic of study, and most research studies do not systematically explore the effects of changing individual problem or solution features on problem solving performance. For example, problems can differ along the following dimensions: physics concepts and principles required to solve it (or a combination of multiple principles), the format of the problem statement (text, picture, diagram, graph), the mathematics required for a solution, values provided for quantities (numeric) or absent (symbolic), presence of additional distracting information, context (e.g., real objects like cars or superficial objects like blocks), and the familiarity of the problem context (e.g., sports compared to nuclear particles).

Areas for Future Study

This review has identified a few prominent gaps in research on problem solving, including research on worked examples, multiple representations, reducing memory load, and adoption of reformed instruction. In particular, the research on the *worked-example effect* in physics is sparse, and there are few guidelines for how to best design instructor solutions. Research on multiple representations is both sparse and contradictory, with little evidence regarding the relationship between use of representations and problem solving performance. Another area that would benefit from future study is developing strategies for effectively reducing memory load while still highlighting important aspects of problem solving. Although there are several instructional strategies and curricula for teaching problem solving, adoption of these practices is not particularly widespread and this warrants additional study.

References

- Adams, W.K., & Wieman, C.E. (2006). Problem solving skill evaluation instrument – Validation studies. In L. McCullough, L. Hsu, & P. Heron (Eds.), *AIP Conference Proceedings Vol. 883: 2006 Physics Education Research Conference* (pp. 18-21). Melville, NY: American Institute of Physics.
- Aguirre, J.M. & Erickson, G. (1984). Students' conceptions about the vector characteristics of three physics concepts. *Journal of Research in Science Teaching*, 21, 437-457.
- Bagno, E., & Eylon, B. (1997). From problem solving to knowledge structure: An example from the domain of electromagnetism. *American Journal of Physics*, 65(8), 726-736.
- Bassok, M., & Holyoak, K.J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(1), 153-166.
- Bing, T.J., & Redish, E.F. (2009). Analyzing problem solving using math in physics: Epistemological framing via warrants. *Physical Review Special Topics – Physics Education Research*, 5(020108).
- Bonham, S.W., Deardorff, D.L., & Beichner, R.J. (2003). Comparison of student performance using web and paper-based homework in college level physics. *Journal of Research in Science Teaching*, 40, 1050-1071.
- Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Cheng, K.K., Thacker, B.A., Cardenas, R.L., & Crouch, C. (2004). Using an online homework system enhances students' learning of physics concepts in an introductory physics course. *American Journal of Physics*, 72(11), 1447-1453.
- Chi, M.T.H. (2006). Two approaches to the study of experts' characteristics. In K.A. Ericsson, N. Charness, R.R. Hoffman, & P.J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). New York, NY: Cambridge University Press.
- Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182.
- Chi, M.T.H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence Vol. 1* (pp. 7-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, E., & Kanim, S.E. (2005). Factors influencing the algebra “reversal error”. *American Journal of Physics*, 73(11), 1072-1078.

Cohen, E., Mason, A., Singh, C., & Yerushalmi, E. (2008). Identifying differences in diagnostic skills between physics students: Students’ self-diagnostic performance given alternative scaffolding. In C. Henderson, M. Sabella, & L. Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference* (pp. 99-102). Melville, NY: American Institute of Physics.

Cui, L., Rebello, N.S., & Bennett, A.G. (2006). College students’ transfer from Calculus to Physics. In P. Heron, L. McCullough, & J. Marx (Eds.), *AIP Conference Proceedings Vol. 818: 2005 Physics Education Research Conference* (pp. 37-40). Melville, NY: American Institute of Physics.

de Jong, T., & Ferguson-Hessler, M.G.M. (1986). Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology*, 78(4), 279-288.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognition and Instruction*, 1, 5-44.

De Leone, C.J., & Gire, E. (2006). Is instructional emphasis on the use of non-mathematical representations worth the effort? In P. Heron, L. McCullough, & J. Marx (Eds.), *AIP Conference Proceedings Vol. 818: 2005 Physics Education Research Conference* (pp. 45-48). Melville, NY: American Institute of Physics.

Dufresne, R.J., Gerace, W.J., Hardiman, P.T., & Mestre, J.P. (1992). Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. *The Journal of the Learning Sciences*, 2(3), 307-331.

Docktor, J. (2009). *Development and validation of a physics problem-solving assessment instrument*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Ferguson-Hessler, M.G.M., & de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor performers. *Cognition and Instruction*, 7(1), 41-54.

Finegold, M., & Mass, R. (1985). Differences in the process of solving physics problems between good problem solvers and poor problem solvers. *Research in Science and Technology Education*, 3, 59-67.

Flores, S., Kanim, S.E., & Kautz, C.H. (2004). Student use of vectors in introductory mechanics. *American Journal of Physics*, 72(4), 460-468.

Gautreau, R., & Novemsky, L. (1997). Concepts first – A small group approach to physics learning. *American Journal of Physics*, 65(5), 418-428.

- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D. (1989). Mechanisms of analogical learning. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning*, (pp. 199-241). London: Cambridge University Press.
- Gertner, A.S., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson, and K. VanLehn (Eds.), *Intelligent Tutoring Systems: 5th International Conference, ITS 2000*, (pp. 133-142). Springer: New York.
- Gick, M.L. (1986). Problem-solving strategies. *Educational Psychologist*, 21(1&2), 99-120.
- Gire, E. (2007). *Between the poles: Locating physics majors in the expert-novice continuum*. Unpublished doctoral dissertation, University of California, San Diego.
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, 68, S59.
- Hammer, D., Elby, A., Scherr, R., & Redish, E.F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89-119). Information Age Publishing.
- Hardiman, P.T., Dufresne, R., & Mestre, J.P. (1989). The relationship between problem categorization and problem solving among experts and novices. *Memory and Cognition*, 17(5), 627-638.
- Harper, K.A. (2001). *Investigating the development of problem solving skills during a freshman physics sequence*. Unpublished doctoral dissertation, The Ohio State University.
- Heller, J.I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction*, 1(2), 177-216.
- Heller, K., & Heller, P.J. (2000). *Competent problem solver – calculus version*. New York: McGraw-Hill.
- Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics*, 60(7), 637-644.
- Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, 60(7), 627-636.
- Henderson, C., & Harper, K.A. (2009). Quiz corrections: Improving learning by encouraging students to reflect on their mistakes. *The Physics Teacher*, 47, 581-586.

- Henderson, C., Yerushalmi, E., Kuo, V., Heller, P., & Heller, K. (2004). Grading student problem solutions: The challenge of sending a consistent message *American Journal of Physics*, 72(2), 164-169.
- Holyoak, K. (1985). The pragmatics of analogical transfer. In G.H. Bower (Ed.), *The psychology of learning and motivation Vol. 19*. San Deigo, CA: Academic Press.
- Holyoak, K.J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15, 332-340.
- Hsu, L., Brewster, E., Foster, T.M., & Harper, K.A. (2004). Resource letter RPS-1: Research in problem solving. *American Journal of Physics*, 72(9), 1147-1156.
- Hsu, L. & Heller, K. (2009). Computer problem-solving coaches. *Proceedings of the National Association for Research in Science Teaching 82nd International Conference*, Garden Grove, CA.
- Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics *Journal of Research on Science Teaching*, 34(6), 551-570.
- Kane, D., & Sherwood, B. (1980). A computer-based course in classical mechanics. *Computers and Education*, 4, 15-36.
- Kashy, E., Sherrill, B.M., Tsai, Y., Thaler, D., Weinshank, D., Engelmann, M., and Morrissey, D.J. (1993). CAPA, an integrated computer assisted personalized assignment system *American Journal of Physics*, 61(12), 1124-1130.
- Kashy, E., Gaff, S.J., Pawley, N.H., Stretch, W.L., Wolfe, S.L., Morrissey, D.J., & Tsai, Y. (1995). Conceptual questions in computer-assisted assignments. *American Journal of Physics*, 63(11), 1000-1005.
- Kohl, P.B., & Finkelstein, N.D. (2005). Student representational competence and self-assessment when solving problems. *Physical Review Special Topics – Physics Education Research*, 1(010104).
- Kohl, P.B., & Finkelstein, N.D. (2008). Patterns of multiple representation use by experts and novices during physics problem solving. *Physical Review Special Topics – Physics Education Research*, 4(010111).
- Kortemeyer, G., Kashy, E., Benenson, W., & Bauer, W. (2008). Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses. *American Journal of Physics*, 76, 438-444.
- Larkin, J. (1979). Processing information for effective problem solving *Engineering Education*, 70(3), 285-288.

Larkin, J. (1981a). Cognition of learning physics. *American Journal of Physics*, 49(6), 534-541.

Larkin, J. (1981b). Enriching formal knowledge: A model for learning to solve textbook physics problems. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 311-344). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner & A.L. Stevens (Eds.), *Mental models*, (pp. 75-98). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science*, 208(4450), 1335-1342.

Lee, Y.-J., Palazzo, D.J., Warnakulassoriya, R., & Pritchard, D.E. (2008). Measuring student learning with item response theory. *Physical Review Special Topics – Physics Education Research*, 4(010102).

Leonard, W.J., Dufresne, R.J., & Mestre, J.P. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems. *American Journal of Physics*, 64(12), 1495-1503.

Maloney, D.P. (1987). Ranking tasks: A new type of test item. *Journal of College Science Teaching*, 16, 510-514.

Maloney, D.P. (1994). Research on problem solving: Physics. In D.L. Gabel (Ed.) *Handbook of research on science teaching and learning* (pp. 327-354). New York: Macmillan.

Mateycik, F., Rebello, N.S., & Jonassen, D. (2010). Students' rating of problem similarity as a measure of problem solving expertise. *Proceedings of the 2010 National Association for Research in Science Teaching (NARST) annual conference*.

Marx, J.D., & Cummings, K. (2010, July). Assessing student's ability to solve textbook-style problems: Update, Part I. Presentation at the summer meeting of the American Association of Physics Teachers, Portland, OR.

Mason, A., & Singh, C. (2009). Revisiting categorization. *Proceedings of the National Association for Research in Science Teaching 82nd International Conference*, Garden Grove, CA.

Meltzer, D.E. (2005). Relation between students' problem-solving performance and representational format. *American Journal of Physics*, 73(5), 463-478.

Meredith, D.C., & Marrongelle, K.A. (2008). How students use mathematical resources in an electrostatics context. *American Journal of Physics*, 76(6), 570-578.

- Mestre, J. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Applied Developmental Psychology, 23*, 9-50.
- Mestre, J.P., Dufresne, R.J., Gerace, W.J., Hardiman, P.T., & Touger, J.S. (1993). Promoting skilled problem-solving behavior among beginning physics students. *Journal of Research in Science Teaching, 30*(3), 303-317.
- Mestre, J., Hart, D.M., Rath, K.A., & Dufresne, R. (2002). The effect of web-based homework on test performance in large enrollment introductory physics courses. *Journal of Computers in Mathematics and Science Teaching, 21*(3), 229-251. Norfolk, VA: AACE.
- Morote, E.-S., & Pritchard, D.E. (2009). What course elements correlate with improvement on tests in introductory Newtonian mechanics? *American Journal of Physics, 77*(8), 746-753.
- Newell, A., Shaw, J.D., & Simon, H.A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151-166.
- Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Nguyen, N.L., & Meltzer, D.E. (2003). Initial understanding of vector concepts among students in introductory physics courses. *American Journal of Physics, 71*(6), 630-638.
- Ogilvie, C.A. (2009). Changes in students' problem-solving strategies in a course that includes context-rich, multifaceted problems. *Physical Review Special Topics – Physics Education Research, 5*(010102).
- O'Kuma, T.L., Maloney, D.P., & Hieggelke, C.J. (2000) *Ranking task exercises in physics*. Upper Saddle River, NJ: Prentice Hall.
- Pascarella, A.M. (2002). *CAPA (Computer-Assisted Personalized Assignments) in a Large University Setting*. Unpublished doctoral dissertation, University of Colorado, Boulder, CO.
- Ploetzner, R., Fehse, E., Kneser, C., & Spada, H. (1999). Learning to relate qualitative and quantitative problem representations in a model-based setting for collaborative problem solving. *The Journal of the Learning Sciences, 8*(2), 177-214.
- Pólya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.
- Reif, F. (1995). *Understanding Basic Mechanics, Text and Workbook* New York: Wiley.
- Reif, F., & Heller, J.I. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist, 17*(2), 102-127.
- Reif, F., & Scott, A. (1999). Teaching scientific thinking skills: Students and computers coaching each other. *American Journal of Physics, 67*, 819-831.

Reeves, L.M., & Weisberg, R.W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115(3), 381-400.

Rosengrant, D., & Mzoughi, T. (2008). Preliminary study of impulse-momentum diagrams. In C. Henderson, M. Sabella, & L. Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference* (pp. 187-190). Melville, NY: American Institute of Physics.

Rosengrant, D., Van Heuvelen, A., & Etkina, E. (2005). Free-body diagrams: Necessary or sufficient? In J. Marx, P. Heron, and S. Franklin (Eds.), *AIP Conference Proceedings Vol. 790: 2004 Physics Education Research Conference* (pp. 177-180). Melville, NY: American Institute of Physics.

Rosengrant, D., Van Heuvelen, A., & Etkina, E. (2006). Case study: Students' use of multiple representations in problem solving. In P. Heron, L. McCullough, & J. Marx (Eds.), *AIP Conference Proceedings Vol. 818: 2005 Physics Education Research Conference* (pp. 49-52). Melville, NY: American Institute of Physics.

Rosengrant, D., Van Heuvelen, A., & Etkina, E. (2009). Do students use and understand free-body diagrams? *Physical Review Special Topics – Physics Education Research*, 5(010108).

Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-639.

Ross, B.H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456-468.

Sabella, M.S., & Redish, E.F. (2007). Knowledge organization and activation in physics problem solving. *American Journal of Physics*, 75(11), 1017-1029.

Sayre, E.C., & Wittmann, M.C. (2008). Plasticity of intermediate mechanics students' coordinate system choice. *Physical Review Special Topics – Physics Education Research*, 4(020105).

Scott, L.A. (2001). *Design and assessment of an interactive physics tutoring environment*. Unpublished doctoral dissertation, University of Pittsburgh.

Shaffer, P.S., & McDermott, L.C. (2005). A research-based approach to improving student understanding of the vector nature of kinematical concepts. *American Journal of Physics*, 73(10), 921-931.

Sherin, B.L. (1996). *The symbolic basis of physical intuition: A study of two symbol systems in physics instruction*. Unpublished doctoral dissertation, University of California, Berkeley.

Sherin, B.L. (2001). How students understand physics equations. *Cognition and Instruction*, 19(4), 479-541.

Sherwood, B.A. (1971). Free-body diagrams (a PLATO lesson). *American Journal of Physics*, 39, 1199-1202.

Simon, H.A. (1978). Information-processing theory of human problem solving. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes*, Vol. 5 (pp. 271-295). Hillsdale, NJ: Erlbaum.

Singh, C. (2002). When physical intuition fails *American Journal of Physics*, 70(11), 1103-1109.

Smith, S.G., & Sherwood, B.A. (1976). Educational uses of the PLATO computer system. *Science*, 192, 344-352.

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

Taconis, R., Ferguson-Hessler, M.G.M., & Broekkamp, H. (2001). Teaching science problem-solving. *Journal of Research in Science Teaching*, 38, 442-468.

Torigoe, E. (2008). *What kind of math matters? A study of the relationship between mathematical ability and success in physics*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Tuminaro, J. (2004). *A cognitive framework for analyzing and describing introductory students use and understanding of mathematics in physics*. Unpublished doctoral dissertation, University of Maryland.

Tuminaro, J., & Redish, E.F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Special Topics – Physics Education Research*, 3(020101).

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

Van Heuvelen, A. (1991b). Overview, case study physics *American Journal of Physics*, 59(10), 898-907.

Van Heuvelen, A. (1994a). *ALPS: Mechanics (Vol. 1)*. Plymouth, MI: Hayden-McNeil Publishing, Inc.

Van Heuvelen, A. (1994b). *ALPS: Electricity and Magnetism (Vol. 2)*. Plymouth, MI: Hayden-McNeil Publishing, Inc.

Van Heuvelen, A. (1995). Experiment problems for mechanics *The Physics Teacher*, 33, 176-180.

- Van Heuvelen, A., Allen, L.D., & Mihas, P. (1999). Experiment problems for electricity and magnetism. *The Physics Teacher*, 37, 482-485.
- Van Heuvelen, A., & Maloney, D. (1999). Playing physics jeopardy. *American Journal of Physics*, 67, 252-256.
- Van Heuvelen, A., & Zou, X. (2001). Multiple representations of work-energy processes. *American Journal of Physics*, 69(2), 184-194.
- Van Weeren, J.H.P., de Mul, F.F.M., Peters, M.J., Kramers-Pals, H., & Roossink, H.J. (1982). Teaching problem-solving in physics: A course in electromagnetism. *American Journal of Physics*, 50, 725-732.
- VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15(3), 147-204.
- Walsh, L.N., Howard, R.G., & Bowe, B. (2007). Phenomenographic study of students' problem solving approaches in physics. *Physical Review Special Topics – Physics Education Research*, 3 (020108).
- Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7(1), 1-39.
- Wright, D.S., & Williams, C.D. (1986). A WISE strategy for introductory physics. *The Physics Teacher*, 24, 211-216.
- Yerushalmi, E., Mason, A., Cohen, E., & Singh, C. (2008). Effects of self diagnosis on subsequent problem solving performance. In C. Henderson, M. Sabella, & L. Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference* (pp. 53-56). Melville, NY: American Institute of Physics.

Curriculum & Instruction in Physics

An abundance of instructional methods and curricular materials have been developed to span all aspects of the standard physics course: lecture, recitation, and laboratories. Some of these instructional reforms aim to make traditional course aspects more engaging or “interactive,” whereas other reforms involve comprehensive changes to the structure of a course, such as combining lecture, recitation, and labs into a single class environment. Advances in technology have also introduced classroom polling technologies into lectures, computers and sensors for collecting and analyzing laboratory data, web-based homework and tutoring systems, and computer animations / simulations of physical phenomena. As a result, the development and evaluation of instructional innovations continues to be an active area of Physics Education Research.

Physicists have been involved in curricular reform for several decades, and one key example is the Physical Sciences Study Committee (PSSC) initiated at MIT in 1956 which produced its first textbook in 1960, along with a laboratory guide, detailed teacher’s guide, and sample assessments (see historical accounts at <http://libraries.mit.edu/archives/exhibits/pssc/> and <http://www.compadre.org/portal/pssc/docs/Haber-Schaim.pdf>). It also included a series of films now available as *Physics Cinema Classics* (Zollman & Fuller, 1994). *PSSC Physics* continued for seven editions (Haber-Schaim, Dodge, Gardner, Shore, & Walter, 1991) and has undergone substantial changes. In the meantime there have been several other PER-based curricular materials developed and tested for their effectiveness, described in this section. There are also several text resources available for physics instructors to learn more about PER-based teaching strategies. Redish’s (2003) book *Teaching Physics with the Physics Suite* gives an overview of several research-based instructional strategies, curricula, and assessments, and also includes a chapter on cognitive science. Books that provide hints for teaching specific physics topics include Arnold Arons’ book *A Guide to Introductory Physics Teaching* (1990) and R. Knight’s *Five Easy Lessons: Strategies for Successful Physics Teaching* (2002).

Research Questions

All examples of instructional reform presented here are (either implicitly or explicitly) related to the research question of evaluating the effectiveness of instructional strategies and materials for teaching physics, and determining the conditions under which the instruction does or does not work. We will, therefore, forego listing specific research questions below since nearly all would be of the “Did it work?” type. Strategies include lecture-based, recitation, and laboratory methods; changes to the overall structure of a class; and general curricular materials such as textbooks and simulations. Many of these examples have been termed “interactive engagement” methods in comparison to traditional, passive lecture methods (Hake, 1998).

Lecture-based methods. These instructional methods seek to enhance the lecture format of a physics course by making in-class experiences interactive through student-student interactions and student-instructor interactions (Melzter & Manivannan, 2002). Reforms to lecture often include opportunities for students to discuss topics with their classmates, like in active learning classes supported by classroom polling technologies (Ding, Reay, Lee, & Bao, 2009; Dufresne, Gerace, Leonard, Mestre, & Wenk, 1996; Mazur, 1997), and *Interactive Lecture Demonstrations* (Sokoloff & Thornton, 2004). Another instructional method is to require students to complete pre-lecture assignments to better prepare them for lecture, such as answering conceptual questions in Just-in-Time Teaching (Novak, Patterson, Gavrinn, & Christian, 1999) or viewing a presentation such as multimedia learning modules (Chen, Stelzer, & Gladding, 2010; Stelzer, Brookes, Gladding, & Mestre, 2010; Stelzer, Gladding, Mestre, & Brookes, 2009). Other lecture-based approaches promote self-reflection about learning, such as requiring students to keep a journal and/or submit weekly reports describing what they’ve learned and what they are still confused about (Etkina, 2000; Harper, Etkina, & Lin, 2003; May & Etkina, 2002).

Recitation or discussion methods. These approaches seek to make traditional, passive recitation sessions more interactive and collaborative. Curricular materials include the *Tutorials in Introductory Physics* developed at the University of Washington (McDermott & Shaffer, 2002) and adaptations thereof, such as *Activity-Based Tutorials* and *Open-Source Tutorials* at the University of Maryland (Elby, 2001; Wittmann, Steinberg, & Redish, 2004a; 2004b). Another recitation method utilizes collaborative learning in discussion sessions, such as a *Cooperative*

Group Problem Solving at the University of Minnesota (Heller & Hollabaugh, 1992; Heller, Keith, & Anderson, 1992).

Laboratory methods. The introduction of technological tools for collecting and analyzing data has transformed the physics laboratory, and spawned several curricula to accompany these changes, including *RealTime Physics* (Sokoloff, Thornton, & Laws, 2004a-d) and video analysis software (Beichner, 1996). Other laboratory materials focus on developing scientific abilities such as *Investigative Science Learning Environments* (Etkina & Van Heuvelen, 2007; Etkina et al., 2010), *Computer-based Problem Solving Labs* (McCullough, 2000) and *Scientific Community Labs* (Lippmann Kung, 2005).

Structural changes to classroom environment. Some methods of instructional reform involve alterations to the standard classroom structure. One example is *Problem-Based Learning*, which modifies the lecture setting to be appropriate for cooperative group work (Duch, Groh, & Allen, 2001). Several other examples revise the course to integrate some aspects of lecture, labs, and recitations into a single setting as in *Physics by Inquiry* (McDermott, 1996a; 1996b), *Workshop Physics* (Laws, 1991; 2004), *Studio Physics* (Cummings, Marx, Thornton, & Kuhl, 1999; Hoellwarth, Moelter, & Knight, 2005; Sorensen, Churukian, Maleki, & Zollman, 2006; Wilson, 1994), *Student-Centered Active Learning Environment for Undergraduate Programs (SCALE-UP)* (Beichner et al., 2007), and *Technology-Enabled Active Learning (TEAL)* (Belcher, 2001; Dori & Belcher, 2005a; Dori et al., 2003).

General instructional strategies and materials This class of reform includes physics textbooks that have been written to incorporate results from physics education research such as *Understanding Physics* (Cummings, Redish, Laws, & Cooney, 2004) and *Matter and Interactions* (Chabay & Sherwood, 2009) and books about teaching physics such as Redish's (2003) *Teaching Physics with the Physics Suite* and Arons' (1991) *A Guide to Introductory Physics Teaching*. It also includes computer-based curricula such as animations and interactive simulations of physical phenomena (Adams et al. 2008a; 2008b; Finkelstein et al., 2005; Steinberg, 2000), and instruction that incorporate computer programming experiences into the course (Chabay & Sherwood, 2008; Redish & Wilson, 1993). Other curricula cited include high

school materials such as *Modeling Instruction* (Halloun & Hestenes, 1985) and *Minds on Physics* (Leonard, Dufresne, Gerace, & Mestre, 1999), since these are sometimes adopted for undergraduate courses for nonscience majors (Brewer, 2008).

Theoretical Frameworks

Most modern theories of instructional design for science courses are based on constructivism and its associated theories, such as situated, sociocultural, ecological, everyday, and distributed cognition (Jonassen & Land, 2000). The constructivist view of teaching and learning emphasizes the active role of a learner takes by interacting with the environment and interpreting information in relation to his/her prior understanding and experiences (Trowbridge, Bybee, & Powell, 2000; von Glasersfeld, 1989). Constructivism has roots in early theories of education, for example the notion that experience and prior knowledge influence learning was discussed by the education theorist Dewey (1938) and apparent in Piaget's theories of cognitive equilibration via assimilation and accommodation. For a historical review of constructivist theories including a discussion of Piaget, see von Glasersfeld (1989).

Instructional strategies and materials for teaching undergraduate physics have incorporated constructivist theories by reforming traditional styles of teaching (passive, lecture-based methods) to be more student-centered and active, approaches collectively referred to as interactive engagement methods (Hake, 1998). Examples of instructional strategies and materials to make lectures more interactive are described in the section on *Lecture-based methods*.

Design principles for creating constructivist learning environments include assigning open-ended tasks (e.g., problems) that are authentic and challenging, giving students opportunities to work collaboratively with their peers, and providing appropriate scaffolding for activities (Jonassen, 1999). Physics curricula that incorporate alternate problems and cooperative group work include *Problem-Based Learning* (Duch et al., 2001) and *Cooperative Group Problem Solving* (Heller & Hollabaugh, 1992; Heller et al., 1992). Other instructional strategies that also utilize teamwork and cooperative learning in the classroom include *Physics by Inquiry* (McDermott 1996a; 1996b), *Tutorials in Introductory Physics* (McDermott & Shaffer, 2002), *Workshop Physics* (Laws 1991; 2004), *Studio Physics* (Cummings et al., 1999; Hoellwarth et al.,

2005; Sorensen et al., 2006; Wilson, 1994), *SCALE-UP* (Beichner et al., 2007), and *TEAL* (Belcher, 2001; Dori & Belcher, 2005a)

Another constructivist-based approach to instruction is the use of *Learning Cycles* (Karplus, 1977; Zollman 1997). A three-phase learning cycle developed for the Science Curriculum Improvement Study (SCIS) included exploration, concept introduction, and concept application (Karplus, 1977). An extension of this work is the 5E cycle of engagement, exploration, explanation, elaboration, and evaluation (Trowbridge, Bybee, & Powell, 2000). Both approaches include an initial “exploratory” period in which students participate in hands-on activities before the formal introduction of concepts. Several physics laboratory curricula incorporate this aspect of early exploration, such as *Physics by Inquiry* (McDermott, 1996a; 1996b) and *Investigative Science Learning Environments* (Etkina & Van Heuvelen, 2007). The upper-division curricula *Paradigms in Physics* also cites the use of learning cycles (Manogue et al., 2001).

Methodology (Data Collection / Sources and Data Analysis)

Contexts. Studies of the effectiveness of particular instructional strategies and/or materials are typically conducted in the context of physics course(s). Oftentimes students’ performance in a reformed course is compared to other sections of a similar course taught traditionally, or to past years in which the course was taught differently. Comparisons are typically made between courses within a single institution, but occasionally researchers make cross-institutional comparisons (Hake, 1998).

Participants. Studies of reformed instruction frequently involve undergraduate students in large-scale introductory courses (algebra-based or calculus-based), and in very few instances there are reforms made to upper-division undergraduate physics courses or graduate student courses. Refinements to instructional methods can take place over a period of several years with multiple cohorts of students.

Data Sources and Analysis The development and evaluation of reformed curricula typically involves a mixture of quantitative and qualitative research methodologies. The development of

curricula is more descriptive and often begins with identifying common student difficulties or misconceptions, so it lends itself to qualitative methods such as interviews with students or faculty. Examples include the process undergone to develop the *Physics by Inquiry* and *Tutorials in Introductory Physics* at the University of Washington (McDermott, 1984; Trowbridge & McDermott, 1980; 1981; Shaffer & McDermott, 1992; 2005) or the procedures to validate clicker questions at The Ohio State University (Ding et al., 2009).

Evaluating the effectiveness of reforms made to large introductory physics courses lends itself to quantitative, statistical methods and the use of quantitative measures. Common data sources for evaluating the effectiveness of instructional strategies or materials include student responses to standard course assignments, such as performance on course exams, laboratories, and homework. Sometimes rubrics are written to facilitate scoring for open-ended items (Etkina et al., 2006b). For instructional reforms that have taken place within the past two decades, a common data source is students' performance on pre-post concept inventories such as the Force Concept Inventory (FCI) and Force and Motion Concept Evaluation (FMCE) or attitudinal surveys. (For a comprehensive discussion of inventories, see the paper section on Assessment). Quantitative data analysis methods are utilized to compare scores, and a commonly cited pre-post test measure is the normalized gain (Hake, 1998). Occasionally qualitative research methods are utilized to examine student discourse and student interactions during a reformed class with the aid of video and audio recording (Dori & Belcher, 2005a).

Findings

Lecture-based Methods

- *Peer discussions and classroom communication systems.* Several methods have been developed to encourage active student involvement during a lecture class, including active learning via classroom polling-facilitated instruction (Dufresne et al., 1996) and *Peer Instruction* (Mazur, 1997). In these approaches, students engage in a questioning cycle that typically includes viewing a question presented by the instructor, discussing ideas with classmates, responding to the question, and engaging in a class-wide discussion. In questionnaires and interviews, students reported having a positive attitude toward active learning using polling technologies and felt that they learned more than they would have in a passive lecture class (Dufresne et al., 1996). Studies of the effectiveness of *Peer Instruction*

cite increased conceptual understanding and problem solving performance (Crouch & Mazur, 2001) and decreased course attrition (Lasry, Mazur, & Watkins, 2008).

The nature and quality of the questions asked is also important (Beatty, Gerace, Leonard, & Dufresne, 2006; Ding et al., 2009). Qualitative questions and questions that require reasoning promote discussion more than questions requiring only a calculation or recall of information (Beatty et al., 2006), and students and instructors both report a preference for such conceptual questions over numerical problems (Dufresne et al., 1996; Keller et al., 2007). Beatty et al. (2006) suggest several ways in which questions can be crafted to include subtle or common student conceptions, such as sequencing two questions to look similar but require different reasoning in an “oops-go-back” format (Beatty et al., 2009). Researchers at The Ohio State University went through a process of validating several clicker questions with a sequence of student interviews, faculty interviews, and classroom response data (Ding et al., 2009). They identified some situations in which students’ perspectives were not anticipated by the question designers, highlighting the importance of validating course materials. The manner in which credit is assigned for responses has been shown to affect the nature of students’ conversations (James, 2006). In high-stakes classrooms group conversations are dominated by high-ability students, whereas low-stakes classrooms promote discussion among students of varying ability levels.

These lecture-based methods frequently employ student response systems such as clickers or other polling/communication devices (see, for example, <http://www.iclicker.com/dnn/> or <http://www.einstruction.com/>). Studies of electronic response systems for gathering student answers indicate that it is not the technology that is important, but the way in which it is used (Caldwell, 2007; Keller et al., 2007; Lasry, 2008). Clickers offer advantages for instructors, such as acquiring an immediate gauge of student understanding and keeping an electronic record of response patterns (Lasry, 2008). For a review of research on clickers and best practices for their use, see Caldwell, 2007.

- *Interactive Lecture Demonstrations (ILDs)*. This curriculum follows a sequence in which students make a prediction about what they expect to see in a physical experiment or demonstration, discuss their prediction with peers, observe the event, and compare the observation to their prediction (Sokoloff & Thornton, 2004). Each stage of the ILD is guided by questions printed on worksheets. Studies evaluating the effectiveness of ILDs indicate

that they significantly improve learning of basic physics concepts measured by the Force and Motion Conceptual Evaluation or FMCE (Sokoloff & Thornton, 1997) and that the prediction phase of classroom demonstrations is particularly important (Crouch et al. 2004).

- *Pre-lecture assignments.* One pre-lecture method, called Just-in-Time teaching (JiTT), requires students to submit answers to conceptual questions electronically before class (Novak et al., 1999). The instructor uses these responses to gauge student understanding and incorporates them into the lecture. JiTT has been shown to improve students' understanding of concepts such as Newton's Third Law (Formica, Easley, & Spraker, 2010). In another pre-lecture method, students view Multimedia Learning Module (MLM) presentations with interspersed questions (<http://research.physics.illinois.edu/per/prelectures.html>) Students who viewed MLMs performed better on assessments of basic content administered both immediately after viewing the MLMs as well as two weeks later compared to students who read the same material in a static text format (Stelzer et al., 2009; 2010) and MLM viewers also performed better on before-lecture assessment questions indicating the presentations improved students' preparation for class (Chen et al., 2010). In addition, when MLMs were implemented students perceived the course as less difficult, they had a more positive attitude toward physics, and they found lectures more valuable since the MLMs allowed class time to be spent on activities to refine their understanding rather than covering basic content (Stelzer et al., 2010).
- *Reflection on learning.* One practice to facilitate students' reflection on their learning is requiring them to keep a journal and/or submit weekly reports (Etkina, 2000; Harper et al., 2003; May & Etkina, 2002). In those studies, students were asked to write down every week what they learned in physics class, how they learned it, and what they still had questions about. Students received credit for submitting responses that were similar in weight to homework. When students' weekly journal statements were coded across several aspects evaluated for their appropriateness (see May & Etkina, 2002 for the code criteria), the researchers found a relationship between performance and reflection quality, where students with high conceptual gains on concept inventories tended to show "reflection on learning that is more articulate and epistemologically sophisticated than students with lower conceptual gains" (May & Etkina, 2002; p. 1249).

Recitation or Discussion Methods

In physics courses, a recitation or discussion session refers to a smaller classroom environment that meets once or twice a week, typically consisting of 15-30 students and taught by a graduate or undergraduate teaching assistant (TA). It is intended to provide students with individualized help and feedback, such as with conceptual activities or problem solving practice. In “traditional” recitations the TA would solve homework problems on the board, whereas reformed sessions are revised to be more interactive.

- *Tutorials*. *Tutorials in Introductory Physics (TIP)* is a supplementary curriculum developed at the University of Washington (McDermott & Shaffer, 2002). It consists of a series of pretests, worksheets, and homework assignments intended to develop conceptual understanding and qualitative reasoning skills. The tutorials are designed for use in a small classroom environment in which students work in groups of 3-4, but it can also be adapted for larger class settings. The *UW Tutorials* were developed from several interview-based studies on student understanding of topics in velocity and acceleration (Rosenquist & McDermott, 1987; Trowbridge & McDermott, 1980; 1981), kinematics graphs (McDermott, Rosenquist, & van Zee, 1987), Newton’s Laws (McDermott, Shaffer, & Somers, 1994), energy and momentum (Lawson & McDermott, 1987; O’Brien Pride, Vokos, & McDermott, 1998), geometric optics (Goldberg & McDermott, 1987; Wosilait, Heron, Shaffer, & McDermott, 1998) and electricity and magnetism (McDermott & Shaffer, 1992; Shaffer & McDermott, 1992). More recent instructional strategies and materials have been developed for topics in light and electromagnetic waves (Ambrose, Heron, Vokos, & McDermott, 1999; Ambrose, Shaffer, Steinberg, & McDermott, 1999; Vokos, Shaffer, Ambrose, & McDermott, 2000; Wosilait, Heron, Shaffer, & McDermott, 1999), gas laws (Kautz, Heron, Loverude, & McDermott, 2005; Kautz, Heron, Shaffer, & McDermott, 2005), thermodynamics (Loverude, Kautz, & Heron, 2002; Cochran & Heron, 2006), hydrostatic pressure (Loverude, Heron, & Kautz, 2010), static equilibrium (Ortiz, Heron, & Shaffer, 2005), and special cases in momentum (Close & Heron, 2010).

The University of Maryland expanded upon the UW tutorial framework to develop *Activity-Based Tutorials (ABT)* which also relate concepts to mathematics and include technological tools such as computers for data acquisition and displaying videos or simulations (Steinberg, Wittmann, & Redish, 1997; Wittmann, Steinberg, & Redish, 2004a;

2004b). Another adaptation from Maryland, called *Open-Source Tutorials (OST)*, permits instructors to modify worksheet materials to meet their needs and places emphasis on developing both students' concepts and epistemologies or beliefs about learning physics (Elby, 2001). Studies of the effectiveness of TIP and ABT tutorials have produced high gains on multiple choice questions and pre-post diagnostic measures such as the Force and Motion Concept Evaluation, but sometimes with mixed attitudinal reactions from students (Finkelstein & Pollock 2005; Mauk & Hingley, 2005; Redish, Saul, & Steinberg, 1997) Although students reported they viewed tutorials as useful and they liked working in groups, very few (20%) said they enjoy tutorials (Finkelstein & Pollock, 2005). A comparison of the three types of tutorials indicates that students using the OST version perform better on course exams and diagnostic tests than TIP or ABT (Smith & Wittmann, 2007).

- *Cooperative learning*. Cooperative learning refers to a strategy for collaborative group activities developed by David and Roger Johnson (D. Johnson, R. Johnson, & Holubec, 1988; D. Johnson, R. Johnson, & Smith, 1991). The principles of cooperative learning have been applied to physics classes in an approach called cooperative group problem solving (Heller & Hollabaugh, 1992; Heller et al., 1992; Hollabaugh, 1995) and also integrated into other approaches, such as Just-In-Time Teaching, Problem-Based Learning (Duch et al., 2001), and Workshop/Studio/SCALE-UP Physics. Cooperative group problem solving necessitates the use of problems appropriate for group work, such as context-rich problems (K. Heller & P. Heller, 2000) or ill-structured problems (Duch et al., 2001) and has been shown to improve problem solving performance (for a comprehensive list of alternate problem types, see the section on Problem Solving Research). However, there are several structural aspects that influence group functioning such as the size of the group (the optimal size is three), composition (heterogeneous by ability), seating arrangement, use of group role assignments, and testing procedures that include both group and individual components (Heller & Hollabaugh, 1992).

Laboratory Methods

These approaches seek to revise traditional labs, which are often described as confirmatory or “cookbook” labs in which students follow step-by-step instructions to verify a result. General approaches to making labs more interactive include engaging students in class discussions before

and at the end of class; revising materials such that students must plan and make decisions; and probing for student understanding during class (see the discussion of laboratory reform in Redish, 2003). Also reviewed here are technological tools used in labs.

- Microcomputer-based labs. With the introduction of microcomputer-based laboratory (MBL) tools such as sonic rangers, it became possible for students to observe a graph produced simultaneously with the motion of an object. Research on these tools and curricula developed to accompany them determined that they improved students' understanding of kinematics concepts and graphs (e.g. position, velocity, and acceleration) in comparison to paper-and-pencil activities (Brasell 1987; Thornton & Sokoloff, 1990). Some researchers attribute the success of MBLs to their real-time data collection, by which students can observe both the motion of an object and a graph of the motion generated simultaneously (Brasell, 1987), whereas others do not see a difference between real-time and delayed-time analysis of motion (Beichner, 1990; Brungardt & Zollman, 1995). A curriculum that utilizes MBL tools to address common student preconceptions is *RealTime Physics Active Learning Laboratories* (Sokoloff et al., 2004a-d) and computer-based problem solving labs at the University of Minnesota (UMN), written in LabVIEW™ (<http://groups.physics.umn.edu/physed/Research/CPSL/MBL.html>). Studies of the *RealTime Physics* laboratories indicate improved student learning of concepts in dynamics, as measured by performance on the FMCE exam (Thornton & Sokoloff, 1997). Studies of the early implementation of UMN computer-based problem solving labs indicated no significant differences between groups using computers and groups using traditional tools with respect to course grades or the TUG-K kinematics inventory, however the experimental (computer) treatment group scored slightly better on the Force Concept Inventory post-test (McCullough, 2000).
- Video analysis software. Another computer-based tool that has been integrated into instruction is video analysis software (Beichner, 1996; Zollman & Fuller, 1994). Students who had opportunities to collect and analyze data resulting from the motion of objects in videos performed better on interpreting kinematics graphs; a teacher demonstration alone was not sufficient for producing significant results (Beichner, 1996). The software was also perceived as useful by students and improved students' comfort with using computers (Escalada & Zollman, 1997).

- *Engaging students in the process of science.* Some instructional strategies for the laboratory seek to engage students in the authentic practice of science in open-ended experiments, such as observing new phenomena, developing hypotheses, designing their own experiments to test them, measurement and uncertainty, and reporting results in written papers (Etkina, Murthy, & Zou, 2006; Reid & Arseneau, 1971; Reif & St. John, 1979). One curriculum that emphasizes the development of scientific abilities is the *Investigative Science Learning Environment (ISLE)* labs (Etkina et al. 2006a; 2010; Van Heuvelen & Etkina, 2006; see also <http://www.islephysics.net>). *ISLE* has been shown to develop students' scientific abilities, such as multiple representations, designing an experiment, collecting and analyzing data, evaluating an experimental outcome, and scientist-like communication (Etkina et al., 2006b; 2010; Karelina & Etkina, 2007). Students who participated in physics laboratory design activities were also able to transfer skills such as evaluating assumptions and experimental uncertainties to a novel task in biology (Etkina et al., 2010). In *ISLE* students go through a learning cycle of observation and exploration, developing explanations, designing experiments to test out and refine their hypotheses, and applying knowledge to solve problems (Etkina et al., 2010; Karplus, 1977). Although labs are a key component of the *ISLE* curriculum, it also includes worksheets and activities appropriate for recitations and large class meetings in the *Physics Active Learning Guide* (Van Heuvelen & Etkina, 2006).

Another laboratory curriculum is the *Problem Solving Labs* developed at the University of Minnesota (see <http://groups.physics.umn.edu/phyped/Research/PSL/pslintro.htm>). Each laboratory experiment is motivated by a context-rich problem, and students are required to complete guiding warm-up questions to reach a prediction before attending the lab session. During lab they discuss predictions in their cooperative groups, and are prompted to make decisions such as planning the data collection and analysis. Although there are research results with respect to UMN's context-rich problems and cooperative group problem solving (Heller & Hollabaugh, 1992; Heller, Keith, & Anderson, 1992; Hollabaugh, 1995), in addition to a comparison of computer and non-computer labs (McCullough, 2000), research on the effectiveness of the problem solving laboratory curricula itself is not currently available and warrants future study.

A laboratory curriculum called *Scientific Community Labs* was developed at the University of Maryland (see <http://umdperg.pbworks.com/Scientific-Community-Lab>). The

labs take a conceptual approach to learning measurement techniques, including the concept of uncertainty (Lippmann Kung, 2005). During lab, students are given a question and must devise an experiment, collect and analyze data, and present their results to the class. Confidence in the results (or “error analysis”) is an explicit part of the discussion. Students using the reformed laboratories exhibited higher gains on a Physics Measurement Questionnaire and more sophisticated reasoning when comparing data sets, such as using “range overlap” rather than percent difference (Lippmann Kung, 2005).

Structural Changes to the Classroom Environment

This section reviews instructional interventions that alter the standard classroom structure. In some instances this includes changing the lecture seating and incorporating technology to accommodate group work on meaningful problems, and in other situations the entire course is restructured to combine lecture, lab, and recitation into a single interactive environment (e.g. workshop and studio class settings).

- *Problem-Based Learning (PBL)*. Developed at the University of Delaware in the early 1990s, Problem-Based Learning has been implemented in a variety of undergraduate science courses, including some in physics (Duch, Groh, & Allen, 2001 <http://www.udel.edu/inst/>). Students work in small cooperative learning teams to solve a complex, real-world problem that may take several weeks to investigate. Lectures are kept to a minimum, and the instructor and/or peer tutors help facilitate the functioning of groups and ask probing questions during class. Research on PBL in physics classes has determined that it helps students develop critical thinking and communication skills (Duch 1996; 1997).
- *Physics by Inquiry (PbI)*. Developed at the University of Washington, the *Physics by Inquiry* curriculum includes a workbook of narratives, laboratory-based experiments and exercises, and supplementary problems (McDermott, 1996a; 1996b see also <http://www.phys.washington.edu/groups/peg/pbi.html>). Students keep a notebook of their observations and exercises, to help them reflect on how their understanding of science concepts changes throughout a course. The targeted audiences for *PbI* include preservice and inservice K-12 science teachers, underprepared students, and nonscience students. Many of the research studies to develop and evaluate the *PbI* materials were conducted concurrently with the studies of *Tutorials in Introductory Physics* described above.

Examples include curricula on light and shadows (Wosilait et al., 1998) or on density (e.g. sinking and floating) (Heron, Loverude, Shaffer, & McDermott, 2003). A study comparing Physics by Inquiry to traditional instruction found that elementary education students in a Pbl course performed significantly better on a quantitative and a qualitative exam problem than physics and engineering students in a traditional introductory physics course (Thacker, Kim, Trefz, & Lea, 1994).

- *Workshop Physics (WP) and Explorations in Physics (EiP)*. *Workshop Physics* was developed at Dickinson College, where the calculus-based physics course was restructured to eliminate formal lectures (Laws, 1991; Laws 2004; http://physics.dickinson.edu/~wp_web/WP_homepage.htm). Instead, students meet for three 2-hour sessions per week to engage in experiments and analyze data using computer tools, such as curve-fitting with spreadsheet software. Research on WP indicates that students prefer the workshop method, have improved performance on conceptual questions (at some but not all institutions), and equal problem solving performance in comparison to traditional sections of the same course (Laws, 1991). The WP activities are published in multiple workbook volumes, called the *Workshop Physics Activity Guide* (Laws, 2004). *Explorations in Physics (EiP)* is the equivalent of *Workshop Physics*, but designed for nonscience students (see http://physics.dickinson.edu/~eip_web/EiP_homepage.htm).
- *Studio Physics*. The term “studio” comes from Jack Wilson’s description of a “Comprehensive Unified Physics Learning Environment” or CUPLE physics studio implemented at Rensselaer (Wilson, 1994). In *Studio Physics*, the lecture and laboratory sessions are combined into a single course session of 30-45 students meeting for two hours, 2-4 times per week (Cummings et al., 1999) with computer-based activities and collaborative group work. Early implementations of *Studio Physics* showed low gains on the Force Concept Inventory exam (similar to scores from traditional courses) but the introduction of research-based methods such as *Interactive Lecture Demonstrations* and *Cooperative Group Problem Solving* significantly improved learning gains (Cummings et al., 1999). Hoellwarth, Moelter, and Knight (2005) found that students in a studio section had significantly higher normalized gain on the Force Concept Inventory and the Force and Motion Conceptual Evaluation than students in a traditional classroom; however students’

scores on quantitative final exam problems were the same or slightly worse in the studio sections.

In another instantiation of studio, called *New Studio Physics*, the lecture remains separate but the recitation and laboratory sessions are combined into a 2-hour session, two times per week (Sorensen et al., 2006). Implementation of *New Studio Physics* found gains on the Force Concept Inventory exam similar to those obtained by other interactive engagement instructional methods.

- *SCALE-UP*. The acronym *SCALE-UP* stands for *Student-Centered Active Learning Environment for Undergraduate Programs*, which was first developed and studied at North Carolina State University in physics classes and has been adopted by several other programs and institutions (Beichner et al. 2007; <http://www.ncsu.edu/per/scaleup.html>). *SCALE-UP* uses a very carefully designed classroom for large enrollment classes (upwards of 100 or more students), in which students work at round tables in teams (typically 3 teams of 3 students at each table) with networked laptops and access to whiteboards. There are multiple display screens around the room. Students engage in hands-on activities including viewing interactive computer simulations, responding to questions or problems, and conducting hypothesis-driven laboratory experiments. Research on *SCALE-UP* cites improved problem solving scores on course exams, conceptual understanding gains as measured by pre-post concept inventories, slightly more positive attitudes, and reduced attrition (see “How do you know it works?” in <http://scaleup.ncsu.edu/FAQs.html>; Beichner et al. 2007; Gaffney, Richards, Kustus, Ding, & Beichner, 2008).
- *TEAL*. One extension of *Studio/SCALE-UP* is the *Technology-Enabled Active Learning (TEAL)* project at the Massachusetts Institute of Technology (Belcher, 2001; Dori et al., 2003). *TEAL* uses the classroom design of *SCALE-UP* which combines lecture, recitation, and laboratories into a single media-enhanced classroom. In addition, it incorporates web-based homework assignments, conceptual questions with clickers, and two- and three-dimensional visualization tools for learning electromagnetism concepts (Dori et al., 2003). Studies of *TEAL* indicate higher learning gains on pre-post tests of multiple-choice conceptual questions (which were adapted from a variety of research-based sources) and a lower failure rate than the traditional course structure (Dori & Belcher, 2005a; 2005b). Open-ended surveys and focus groups indicated that students’ attitudes toward the *TEAL*

classrooms varied, with several students expressing positive comments regarding the usefulness of the collaborative activities and others expressing they preferred a lecture format, or thought that group work was too unguided (Dori & Belcher, 2005a).

General Instructional Strategies and Materials

- *Textbooks.* *Understanding Physics* is a text written by Cummings, Laws, Redish, & Cooney (2004) based on results from Physics Education Research. The text includes touchstone example problems and alternate types of problems, microcomputer-based laboratory tools, and topic organization to facilitate student understanding. The text *Physics for Scientists and Engineers: A Strategic Approach* by Knight (2004) also incorporates PER-based ideas such as multiple representations, alternate problem types, and a four-step problem solving strategy for all worked examples: model, visualize, solve and assess. The curriculum *Six Ideas that Shaped Physics* was developed and tested over a period of more than eight years as part of the Introductory University Physics Project (IUPP) at Pomona College (<http://www.physics.pomona.edu/sixideas/>; Moore, 2003). The volumes are structured to emphasize the hierarchical structure of physics (e.g., Volume C on conservation laws) and provide support for active learning. Courses using the materials found very high gains on conceptual inventories including the Force Concept Inventory (FCI) and Brief Electricity & Magnetism Assessment (BEMA), with normalized gains in the range of 0.50-0.70 (see “Evidence for Success”, <http://www.physics.pomona.edu/sixideas/sisuc.html>).

Another research-based curricula is *Matter & Interactions (M&I)* developed at North Carolina State University (Chabay & Sherwood, 2009; <http://matterandinteractions.org/>). *M&I* approaches calculus-based introductory physics from a “modern” perspective highlighting the atomic nature of matter, relativity, and computational modeling of complex systems (Chabay & Sherwood, 1999; 2004; 2006). *Matter & Interactions* has resulted in higher gains on the Brief Electricity & Magnetism Assessment (BEMA) pre-post diagnostic test (Kohlmyer et al., 2009).

- *Problem solving.* For a review of instructional strategies and materials for teaching problem solving, see the section on Problem Solving Research. The methods reviewed there include alternate types of problems, qualitative approaches, teaching an explicit framework, and web-based homework and computer tutors for problem solving.

- *Animations, simulations, and computer programming.* A recent trend in instructional strategies and materials is the use of technological tools in the classroom, such as animations, simulations, visualization tools, and computer programming experiences. With respect to research on animations, Dancy and Beichner (2006) compared students' responses to Force Concept Inventory questions with static pictures and computer animations for the same questions. They found that the students were better able to interpret the question's intention when viewed in animated form, and they gave an answer more diagnostic of their understanding. Another study by Yeo, Loss, Zadnik, Harrison, & Treagust (2004) found that students using interactive media often retained their intuitive (incorrect) conceptions about motion unless they had explicit guidance from researcher interventions. Steinberg (2000) did not find performance differences between students who interacted with a simulation for learning air resistance compared to those who completed paper-and-pencil versions, however students who interacted with computer simulations were more reliant upon the computer to provide them with the correct answer. Computer visualization tools have been shown to be particularly effective for abstract concepts in electricity and magnetism, such as the concept of a field (Casperson & Linn, 2006; Dori & Belcher, 2005a; 2005b; Stelzer et al., 2009; 2010). Although textbooks often include access to online visualization or animation tools, research support for their development and effectiveness is not always available and/or cited (see for example ActivPhysics OnLine <http://www.aw-bc.com/knight>).

One of the most widely cited projects to develop and test interactive science simulations is the University of Colorado's PhET project (<http://phet.colorado.edu>; Perkins et al., 2006). The web site includes a range of free simulations of phenomena for physics, biology, chemistry, earth science, and math; it also includes a bank of activities created by teachers to accompany the simulations. The physics section includes (to date) more than 70 simulations. There have been several research studies published regarding the effectiveness of the PhET simulations (<http://phet.colorado.edu/en/research>; Adams et al., 2008a; 2008b; McKagan et al., 2008). One study in particular (Finkelstein et al., 2005) found that students who used a computer simulation to explore direct current circuits rather than real light bulbs, meters, and wires had higher scores on their lab report, scored higher on final exam questions pertaining to circuits, and were more successful at performing tasks including assembling a

real circuit. They cite that computer simulations are particularly useful for instruction in that they allow students to view “hidden” phenomena, such as current flow, and focus their attention to relevant details of the representation.

Finally, a few studies have explored the use of computer programming in physics courses. An early study with programming in a course for physics majors was the project *M.U.P.P.E.T.: Maryland University Project in Physics and Educational Technology* (Redish & Wilson, 1993). The *M.U.P.P.E.T.* project determined that programming in introductory physics facilitated the use of more realistic problems and modern topics, and the use of carefully designed projects could encourage students’ use of qualitative and analytic reasoning skills. Another instantiation of computer programming in introductory physics is the use of VPython in the *Matter & Interactions* curriculum (Chabay & Sherwood, 2008). In addition to teaching students some basic programming and debugging skills, the computational tools also facilitates students’ exploration of complex models and understanding of vectors. For a more comprehensive review of research on integrating computational physics into the undergraduate curriculum, see the April 2008 theme issue of the *American Journal of Physics* (volume 76, issue 4).

- *Homework.* There are few studies related to effective procedures for the assignment, collection, and grading of homework in large enrollment courses. One study found that allowing students some choice in which problems to solve and making a fraction of problem solutions available before homework is due provided more timely feedback to students, and resulted in an increase in self-direction (Bao, Stonebraker, & Sadaghiani, 2008).
- *Preservice and inservice teachers.* As mentioned previously, the *Physics by Inquiry* laboratory materials are appropriate for educating preservice and inservice K-12 teachers about physical science (McDermott, 1996a; 1996b). A different curriculum targeted toward elementary teachers is *Physics and Everyday Thinking (PET)* (see an overview of the research project at <http://petproject.sdsu.edu/> and sample materials at <http://www.petpset.net/>). PET has shown a shift toward expert-like attitudes as measured by the Colorado Learning Attitudes about Science Survey (Otero & Gray, 2008).
- *High school curricula.* Although not reviewed extensively here, it is worth mentioning that there have been several texts and workbooks developed for use in middle school physical science and high school physics, which are sometimes adapted for undergraduate courses

with nonscience majors. Examples include *Modeling Instruction* (see <http://modeling.asu.edu/>; Halloun & Hestenes, 1985; Brewster, 2008), *Minds on Physics* (Leonard et al. 1999), *It's About Time: Active Physics* (see <http://www.its-about-time.com/htmls/ap.html> and <http://www.iactivephysics.net/>), *Tools for Scientific Thinking* (Thornton, 1989; <http://www.vernier.com/cmat/tst.html>), and a curriculum for 7th-8th graders called *InterActions in Physical Science* (see <http://www.interactionsinfo.net/>).

- *Upper-division curricula*. One curriculum that reforms the junior-level and senior-level courses for physics majors is the *Paradigms in Physics* project at Oregon State University (Manogue et al., 2001; McIntyre, Tate, & Manogue, 2008). It includes revisions to the structure or order of topics, content of courses, and instructional methods, in an effort to better reflect physicists' views and experiences in the field. Examples include adding laboratory experiments as a component in several courses, making a stronger link between mathematics and physics, and introducing computational examples and problems (McIntyre et al., 2008). Evaluation of the program indicates improved retention of physics majors and improved physics problem solving abilities (Manogue et al., 2001). The development and evaluation of *UW Tutorials* curricula for upper-division mechanics courses also shows promising effects on learning advanced topics (Ambrose, 2004). The University of Colorado-Boulder has transformed their upper-division electricity and magnetism course (Chasteen & Pollock, 2009; Pollock, Chasteen, Dubson, & Perkins, 2010) and quantum mechanics course (Goldhaber, Pollock, Dubson, Beale, & Perkins, 2009) by integrating active engagement methods such as using conceptual questions during lecture, tutorials during recitations, and homework that utilizes multiple representations. They have identified several areas in which upper-division students have difficulty, such as communicating understanding, visualizing fields, problem solving techniques (Chasteen & Pollock, 2009) and the use of conceptual clicker questions can give faculty insight into students' understanding (Pollock et al., 2010).

Reform to graduate education is less common. Carr & McKagan (2009) researched the implementation of PER methods into a graduate quantum mechanics course, including changes to the content, textbook, instructional methods, and assessment. Materials were adapted during the course in response to student surveys and interviews. They found that students became more interested in course material when the instructor made explicit

connections to modern research, and students were responsive and positive toward the lecture modifications, such as the increase in peer discussions during class.

Although many of the instructional materials described here are stand-alone curricula, they are sometimes implemented in combination. Redish's (2003) text outlines components of the *Physics Suite*, which integrates several curricula into an activity-based physics program: the *Understanding Physics* text (Cummings et al., 2004), alternate types of problems, *Interactive Lecture Demonstrations* (Sokoloff & Thornton, 2004), *Workshop Physics* (Laws, 2004), and computer tools in workshops and labs with *RealTime Physics* (Sokoloff et al., 2004a-d).

Strengths & Limitations

Strengths. A strength of research on curriculum and instruction in physics is the sheer abundance of methods that have been developed and evaluated with research. In particular, many early approaches have been tested and refined over several decades to the point of becoming mainstream in introductory undergraduate courses (such as the use of conceptual questions and polling technologies, like clickers). Adoption of research-based instruction is also spreading as a result of New Faculty Workshops sponsored by the American Association of Physics Teachers, American Physical Society, and American Astronomical Society (<http://www.aapt.org/Conferences/newfaculty/nfw.cfm>).

Limitations. This review identified some instructional strategies and curricula that have not been extensively tested for their effectiveness in physics, such as the *Just-in-Time* teaching method and several laboratory-based curricula. In addition, although some textbooks are PER-based, many of the most commonly-used texts for undergraduate courses are not, and this can create a mismatch between the goals of reformed instruction and the materials actually used in the course.

Areas for Future Study

- *Upper-division undergraduate courses and graduate education.* An area of future study includes expanding instructional strategies and materials that have been developed for introductory undergraduate physics courses to other populations, such as upper-division

(junior- and senior-level) courses for physics majors and courses for graduate students.

Although there have been some studies of curriculum and instruction in upper-division (Carr & McKagan, 2009; Manogue et al., 2001; Pollock et al., 2010) this area would benefit from additional research.

- *High School - University cross-curricular studies.* Although several of the physics curricula are identified as appropriate for a particular student population, it is possible that some curricula for high school might be appropriate for university non-science majors (see Brewster, 2008) and that university curricula might be appropriate for advanced high school courses. The application of instructional strategies and materials to alternate groups of students is a possible area for future research.
- *Courses for biology and pre-medicine students.* There is some current interest in reforming the introductory course for biology and pre-medicine students (Redish & Hammer, 2009). Additional discussion is needed to determine how current courses are (or are not) meeting the needs of this student population, and research is needed to develop and test reformed instruction and curricula. At least one research group in Germany has developed and tested the effectiveness of a reformed physics laboratory curriculum for medical students (Plomer, Jessen, Rangelov, & Meyer, 2010). They found that including medical applications (e.g., connecting electricity and neural functions) improved students' attitudes towards physics and improved their ability to relate concepts of physics and medicine as measured by the construction of concept maps.
- *The influence of instructor modifications on the success of reformed curricula.* Instructors do not always implement PER-based instructional methods in the way that they were intended by the developers, and it is unclear what effect this has on the effectiveness of the method. An area of current research is to explore ways in which instructors modify methods when implementing them in their classes, their reasoning for doing so, and how this influences student learning. For examples of some studies, see the section on Attitudes and Beliefs about Teaching and Learning Physics.
- *Lab instruction.* Much of the early research on labs focused on the integration of computer-based sensors and technological tools into the laboratory, but it is possible that many institutions still use confirmatory-style, "cookbook" labs, despite the use of reformed instruction in other aspects of the course (such as lecture and recitation). With the exceptions

of ISLE, RealTime Physics and integrated course structures (workshop, studio, SCALE-UP, and TEAL), there has not been much published research on the effectiveness of laboratory curricula.

- Textbooks. Although there is unpublished research about the “readability” of introductory physics textbooks (Campbell, 2008), there is a need for more research on how students read and interact with physics textbooks and components in differing formats, such as equations, diagrams, text, colored photos, etc. Cognitive load theory suggests that many “enhancements” to textbooks may create extraneous memory load that detract from, rather than add to learning, although little research exists on the design of effective textbooks (see Plass, Moreno & Brunken, 2010, and Sweller, forthcoming for reviews).
- Instructional technology. Although there were several aspects of technology identified here (e.g. classroom polling technologies, multimedia presentations, computer data collection and analysis tools, and computer programming) it is expected that technology and computer-based instruction will continue to evolve, and this will require future research on the effective use of these technologies in classrooms.

References

Adams, W.K., Reid, S., LeMaster, R., McKagan, S.B., Perkins, K.K., Dubson, M., & Wieman, C.E. (2008a). A study of educational simulations part I: Engagement and learning *Journal of Interactive Learning Research*, 19(3), 397-419.

Adams, W.K., Reid, S., LeMaster, R., McKagan, S.B., Perkins, K.K., Dubson, M., & Wieman, C.E. (2008b). A study of educational simulations part II: Interface design *Journal of Interactive Learning Research*, 19(4), 551-577.

Ambrose, B.S. (2004). Investigating student understanding in intermediate mechanics: Identifying the need for a tutorial approach to instruction. *American Journal of Physics*, 72(4), 453-459.

Ambrose, B.S., Heron, P.R.L., Vokos, S., & McDermott, L.C. (1999). Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena *American Journal of Physics*, 67(10), 891-898.

Ambrose, B.S., Shaffer, P.S., Steinberg, R.N., & McDermott, L.C. (1999). An investigation of student understanding of single-slit diffraction and double-slit interference *American Journal of Physics*, 67(2), 146-155.

- Arons, A.B. (1990). *A guide to introductory physics teaching*. New York: John Wiley & Sons, Inc.
- Bao, L., Stonebraker, S.R., & Sadaghiani, H. (2008). A flexible homework method. *American Journal of Physics*, 76(9), 878-881.
- Beatty, I.D., Gerace, W.J., Leonard, W.J., & Dufresne, R.J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, 74(1), 31-39.
- Belcher, J. (2001). Studio physics at MIT. *MIT Physics Annual 2001*, 58-64. Retrieved October 6, 2010 from <http://web.mit.edu/kzhang/MacData/afs.course/other/viz/EM/visualizations/resources/PhysicsNewsLetter.pdf>
- Beichner, R. (1990). The effect of simultaneous motion presentation and graph generation in a kinematics lab. *Journal of Research in Science Teaching*, 27(8), 803-815.
- Beichner, R.J. (1996). The impact of video motion analysis on kinematics graph interpretation skills. *American Journal of Physics*, 64(10), 1272-1277.
- Beichner, R.J., Saul, J.M., Abbott, D.S., Morse, J.J., Deardorff, D.L., Allain, R.J., Bonham, S.W., Dancy, M.H., & Risley, J.S. (2007). The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. In E.F. Redish and P.J. Cooney (Eds.), *Reviews in PER volume 1: Research-based reform of university physics*. College Park, MD: American Association of Physics Teachers.
- Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, 24(4), 385-395.
- Brewe, E. (2008). Modeling theory applied: Modeling Instruction in introductory physics. *American Journal of Physics*, 76(12), 1155-1160.
- Brungardt, J., & Zollman, D. (1995). Influence of interactive videodisc instruction using simultaneous-time analysis on kinematics graphing skills of high school physics students. *Journal of Research in Science Teaching*, 32(8), 855-869.
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE Life Sciences Education*, 6(1), 9-20.
- Campbell, J.L. (2008). A quantitative comparison of the composition and linguistic properties of introductory physics textbooks. Unpublished master's dissertation, University of Arkansas.
- Carr, L.D., & McKagan, S.B. (2009). Graduate quantum mechanics reform. *American Journal of Physics*, 77(4), 308-319.

Casperson, J.M., & Linn, M.C. (2006). Using visualizations to teach electrostatics. *American Journal of Physics*, 74(4), 316-323.

Chabay, R.W., & Sherwood, B.A. (1999). Bringing atoms into first-year physics. *American Journal of Physics*, 67(12), 1045-1050.

Chabay, R., & Sherwood, B. (2004). Modern mechanics. *American Journal of Physics*, 72(4), 439-445.

Chabay, R., & Sherwood, B. (2006). Restructuring the introductory electricity and magnetism course. *American Journal of Physics*, 74(4), 329-336.

Chabay, R., & Sherwood, B. (2008). Computational physics in the introductory calculus-based course. *American Journal of Physics*, 76(4), 307-313.

Chabay, R.W., & Sherwood, B.A. (2009). *Matter and interactions (3rd Ed.)*. Hoboken, NJ: John Wiley & Sons, Inc.

Chasteen, S.V., & Pollock, S.J. (2009). A research-based approach to assessing student learning issues in upper-division electricity and magnetism. In M. Sabella, C. Henderson, and C. Singh (Eds.), *AIP Conference Proceedings Vol. 1179: 2009 Physics Education Research Conference* (pp. 7-10). Melville, NY: American Institute of Physics.

Chen, Z., Stelzer, T., & Gladding, G. (2010). Using multimedia modules to better prepare students for introductory physics lecture. *Physical Review Special Topics – Physics Education Research*, 6(010108).

Close, H.G., & Heron, P.R.L. (2010). Research as a guide for improving student learning: An example from momentum conservation. *American Journal of Physics*, 78(9), 961-969.

Cochran, M.J., & Heron, P.R.L. (2006). Development and assessment of research-based tutorials on heat engines and the second law of thermodynamics. *American Journal of Physics*, 74(8), 734-741.

Crouch, C.H., Fagen, A.P., Callan, J.P., & Mazur, E. (2004). Classroom demonstrations: Learning tools or entertainment? *American Journal of Physics*, 72(6), 835-838.

Crouch, C., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970-977.

Cummings, K., Laws, P., Redish, E., & Cooney, P. (2004). *Understanding physics*. Hoboken, NJ: John Wiley & Sons, Inc.

Cummings, K., Marx, J., Thornton, R., & Kuhl, D. (1999). Evaluating innovation in studio physics. *American Journal of Physics, Physics Education Research Supplement*, 67(7), S38-S44.

- Dancy, M.H., & Beichner, R. (2006). Impact of animation on assessment of conceptual understanding in physics. *Physical Review Special Topics – Physics Education Research*, 2(010104).
- Dewey, J. (1938/1998). *Experience and education* (60th Anniversary Edition). Indianapolis, Indiana: Kappa Delta Pi.
- Ding, L., Reay, N.W., Lee, N., & Bao, L. (2009). Are we asking the right questions? Validating clicker question sequences by student interviews. *American Journal of Physics*, 77(7), 643-650.
- Dori, Y.J., Belcher, J., Bessette, M., Danziger, M., McKinney, A., & Hult, E. (2003). Technology for active learning. *Materials Today*, 6(12), 44-49
- Dori, Y.J., & Belcher, J. (2005a). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *The Journal of the Learning Sciences*, 14(2), 243-279.
- Dori, Y.J., & Belcher, J. (2005b). Learning electromagnetism with visualizations and active learning. In J.K. Gilbert (Ed.), *Visualization in science education* (pp. 187-216). The Netherlands: Springer.
- Duch, B.J. (1996). Problem-based learning in physics: The power of students teaching students. *Journal of College Science Teaching*, 15(5), 326-329.
- Duch, B.J. (1997). Problem-based learning in physics: Making connections with the real world. In E.F. Redish & J.S. Rigden (Eds.), *AIP Conference Proceedings Vol. 399: The changing role of physics departments in modern universities, proceedings of the International Conference on Undergraduate Physics Education* (pp. 557-565). Melville, NY: American Institute of Physics.
- Duch, B.J., Groh, S.E., & Allen, D.E. (2001). *The power of problem-based learning*. Sterling, VA: Stylus Publishing, LLC.
- Dufresne, R.J., Gerace, W.J., Leonard, W.J., Mestre, J.P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, 7(2), 3-47.
- Elby, A. (2001). Helping physics students learn how to learn. *American Journal of Physics, Physics Education Supplement*, 69(7), S54-S64.
- Escalada, L.T., & Zollman, D.A. (1997). An investigation of the effects of using interactive digital video in a physics classroom on student learning and attitudes. *Journal of Research in Science Teaching*, 34(5), 467-489.
- Etkina, E. (2000). Weekly reports: A two-way feedback tool. *Science Education*, 84(5), 594-605.

Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., Hmelo-Silver, C.E. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. *Journal of the Learning Sciences*, 19(1), 54-98.

Etkina, E., Murthy, S., & Zou, X. (2006a). Using introductory labs to engage students in experimental design. *American Journal of Physics*, 74(11), 979-986.

Etkina, E., & Van Heuvelen, A. (2007). Investigative Science Learning Environment – A science process approach to learning physics. In E.F. Redish and P.J. Cooney (Eds.) *Reviews in PER volume 1: Research-based reform of university physics*. College Park, MD: American Association of Physics Teachers.

Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D.T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006b). Scientific abilities and their assessment. *Physical Review Special Topics – Physics Education Research*, 2(020103).

Finkelstein, N.D., Adams, W.K., Keller, C.J., Kohl, P.B., Perkins, K.K., Podolefsky, N.S., Reid, S., & LeMaster, R. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics – Physics Education Research*, 1(010103).

Finkelstein, N.D., & Pollock, S.J. (2005). Replicating and understanding successful innovations: Implementing tutorials in introductory physics. *Physical Review Special Topics – Physics Education Research*, 1 (010101).

Formica, S.P., Easley, J.L., & Spraker, M.C. (2010). Transforming common-sense beliefs into Newtonian thinking through Just-In-Time Teaching. *Physical Review Special Topics – Physics Education Research*, 6(020106).

Gaffney, J., Richards, E. Kustus, M.B., Ding, L. & Beichner, R. (2008). Scaling up education reform. *Journal of College Science Teaching*, 37(5), 48-53.

Goldberg, F.M., & McDermott, L.C. (1987). An investigation of student understanding of the real image formed by a converging lens or concave mirror. *American Journal of Physics*, 55(2), 108-119.

Goldhaber, S., Pollock, S., Dubson, M., Beale, P., & Perkins, K. (2009). Transforming upper-division quantum mechanics: Learning goals and assessment. In M. Sabella, C. Henderson, and C. Singh (Eds.), *AIP Conference Proceedings Vol. 1179: 2009 Physics Education Research Conference* (pp. 145-148). Melville, NY: American Institute of Physics.

Haber-Schaim, U., Dodge, J.H., Gardner, R., Shore, E.A., & Walter, F. (1991). *PSSC Physics (7th Ed.)*. Dubuque, IA: Kendall Hunt Publishing Company.

- Hake, R.R. (1998). Interactive-engagement versus traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74.
- Halloun, I., & Hestenes, D. (1987). Modeling instruction in mechanics. *American Journal of Physics*, 55(5), 455-462.
- Harper, K.A., Etkina, E., & Lin, Y. (2003). Encouraging and analyzing student questions in a large physics course: Meaningful patterns for instructors. *Journal of Research in Science Teaching*, 40(8), 776-791.
- Heller, K., & Heller, P. (2000). *Competent problem solver – calculus version*. New York: McGraw-Hill.
- Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics*, 60(7), 637-644.
- Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, 60(7), 627-636.
- Heron, P.R.L., Loverude, M.E., Shaffer, P.S., & McDermott, L.C. (2003). Helping students develop an understanding of Archimedes' principle. II. Development of research-based instructional materials. *American Journal of Physics*, 71(11), 1188-1195.
- Hoellwarth, C., Moelter, M., & Knight, R. (2005). A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms. *American Journal of Physics*, 74, 374.
- Hollabaugh, M. (1995). *Physics problem solving in cooperative learning groups*. Unpublished doctoral dissertation, University of Minnesota.
- James, M.C. (2006). The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, 74(8), 689-691.
- Johnson, D.W., Johnson, R.T., & Holubec, E.J. (1988). *Cooperation in the classroom*. Edina, MN: Interaction Book Company.
- Johnson, D.W., Johnson, R.T., & Smith, K.A. (1991). *Active learning: Cooperation in the college classroom*. Edina, MN: Interaction Book Company.
- Jonassen, D. (1999). Designing constructivist learning environments. In C.M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory volume II* (pp. 217-240). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Jonassen, D.H., & Land, S.M. (2000). *Theoretical foundations of learning environments*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Karelina, A., & Etkina, E. (2007). Acting like a physicist: Student approach study to experimental design. *Physical Review Special Topics – Physics Education Research*, 3(020106).
- Karplus, R. (1977). Science teaching and the development of reasoning. *Journal of Research in Science Teaching*, 14(2), 169-175.
- Kautz, C.H., Heron, P.R.L., Loverude, M.E., & McDermott, L.C. (2005). Student understanding of the ideal gas law, Part I: A macroscopic perspective. *American Journal of Physics*, 73(11), 1055-1063.
- Kautz, C.H., Heron, P.R.L., Shaffer, P.S., & McDermott, L.C. (2005). Student understanding of the ideal gas law, Part II: A microscopic perspective. *American Journal of Physics*, 73(11), 1064-1071.
- Keller, C., Finkelstein, N.D., Perkins, K., Pollock, S.J., Turpen, C., & Dubson, M. (2007). Research-based practices for effective clicker use. In L. Hsu, C. Henderson, & L. McCullough (Eds.), *AIP Conference Proceedings Vol. 951: 2007 Physics Education Research Conference* (pp. 128-131). Melville, NY: American Institute of Physics.
- Knight, R.D. (2002). *Five easy lessons: strategies for successful physics teaching*. Glenview, IL: Addison Wesley.
- Knight, R.D. (2004). *Physics for scientists and engineers: A strategic approach with modern physics and MasteringPhysics*. San Francisco, CA: Addison Wesley.
- Kohlmyer, M.A., Caballero, M.D., Catrambone, R., Chabay, R.W., Ding, L., Haugan, M.P., Marr, M.J., Sherwood, B.A., & Schatz, M.F. (2009). Tale of two curricula: The performance of 2000 students in introductory electromagnetism. *Physical Review Special Topics – Physics Education Research*, 5(020105).
- Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher*, 46(4), 242-244.
- Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics*, 76(11), 1066-1069.
- Laws, P. (1991). Calculus-based physics without lectures. *Physics Today*, 44(12), 24-31.
- Laws, P.W. (2004). *Workshop Physics activity guide (4 volumes)*. John Wiley & Sons, Inc.
- Lawson, R.A., & McDermott, L.C. (1987). Student understanding of the work-energy and impulse-momentum theorems. *American Journal of Physics*, 55(9), 811-817.

Leonard, W.J., Dufresne, R.J., Gerace, W.J., & Mestre, J.P. (1999-2000). *Minds-On Physics: Activities and reader* (6 volumes). Dubuque, IA: Kendall/Hunt Publishing Company.

Lippmann Kung, R. (2005). Teaching the concepts of measurement: An example of a concept-based laboratory course. *American Journal of Physics*, 73(8), 771-777.

Loverude, M.E., Heron, P.R.L., & Kautz, C.H. (2010). Identifying and addressing student difficulties with hydrostatic pressure. *American Journal of Physics*, 78(1), 75-85.

Loverude, M.E., Kautz, C.H., & Heron, P.R.L. (2002). Student understanding of the first law of thermodynamics: Relating work to the adiabatic compression of an ideal gas. *American Journal of Physics*, 70(2), 137-148.

Manogue, C.A., Siemens, P.J., Tate, J., Browne, K., Niess, M.L., & Wolfer, A.J. (2001). Paradigms in Physics: A new upper-division curriculum. *American Journal of Physics*, 69(9), 978-990.

Mauk, H.V., & Hingley, D. (2005). Student understanding of induced current: Using tutorials in introductory physics to teach electricity and magnetism. *American Journal of Physics*, 73(12), 1164-1171.

May, D.B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics*, 70(12), 1249-1258.

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

McCullough, L. (2000). *The effect of introducing computers into an introductory problem-solving laboratory*. Unpublished doctoral dissertation, University of Minnesota.

McDermott, L.C. (1984). Research on conceptual understanding in mechanics. *Physics Today*, 37(7), 24-32.

McDermott, L.C., & The Physics Education Group at the University of Washington (1996a). *Physics by Inquiry Volume I*. New York: John Wiley & Sons, Inc.

McDermott, L.C., & The Physics Education Group at the University of Washington (1996b). *Physics by Inquiry Volume II*. New York: John Wiley & Sons, Inc.

McDermott, L.C., Rosenquist, M.L., & van Zee, E.H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55(6), 503-513.

McDermott, L.C., & Shaffer, P.S. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics*, 60(11), 994-1003.

McDermott, L.C., Shaffer, P.S., & Somers, M.D. (1994). Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood's machine *American Journal of Physics*, 62(1), 46-55.

McDermott, L.C., Shaffer, P.S., & The Physics Education Group at the University of Washington (2002). *Tutorials in Introductory Physics*. Upper Saddle River, NJ: Prentice-Hall, Inc.

McIntyre, D.H., Tate, J., & Manogue, C.A. (2008). Integrating computational activities into the upper-level Paradigms in Physics curriculum at Oregon State University *American Journal of Physics*, 76(4&5), 340-346.

McKagan, S.B., Perkins, K.K., Dubson, M., Malley, C., Reid, S., LeMaster, R., & Wieman, C.E. (2008). Developing and researching PhET simulations for teaching quantum mechanics. *American Journal of Physics*, 76(4), 406-417.

Meltzer, D.E., & Manivannan, K. (2002). Transforming the lecture-hall environment: The fully interactive physics lecture. *American Journal of Physics*, 70(6), 639-654.

Moore, T.A. (2003). *Six ideas that shaped physics* (6 volumes: C, N, R, E, Q, T). McGraw-Hill.

Novak, G.M., Patterson, E.T., Gavrin, A.D., & Christian, W. (1999). *Just in time teaching*. Upper Saddle River, NJ: Prentice Hall.

O'Brien Pride, T., Vokos, S., & McDermott, L.C. (1998). The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems. *American Journal of Physics*, 66(2), 147-157.

Ortiz, L.G., Heron, P.R.L., & Shaffer, P.S. (2005). Student understanding of static equilibrium: Predicting and accounting for balancing. *American Journal of Physics*, 73(6), 545-553.

Otero, V.K., & Gray, K.E. (2008). Attitudinal gains across multiple universities using the Physics and Everyday Thinking curriculum *Physical Review Special Topics – Physics Education Research*, 4(020104).

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics *The Physics Teacher*, 44(1), 18-23.

Plass, J.L., Moreno, R., & Brunken, R. (Eds.) (2010). *Cognitive Load Theory*. New York, NY: Cambridge University Press.

Plomer, M., Jessen, K., Rangelov, G., & Meyer, M. (2010). Teaching physics in a physiologically meaningful manner *Physical Review Special Topics - Physics Education Research*, 6(020116).

- Pollock, S.J., Chasteen, S.V., Dubson, M., & Perkins, K.K. (2010). The use of concept tests and peer instruction in upper-division physics. In C. Singh, M. Sabella, and S. Rebello (Eds.) *AIP Conference Proceedings Vol. 1289: 2010 Physics Education Research Conference* (pp. 261-264). Melville, NY: American Institute of Physics.
- Redish, E.F. (2003). *Teaching physics with the physics suite*. Hoboken, NJ: John Wiley & Sons, Inc.
- Redish, E.F., & Hammer, D. (2009). Reinventing college physics for biologists: Explicating an epistemological curriculum. *American Journal of Physics*, 77(7), 629-642.
- Redish, E.F., Saul, J.M., & Steinberg, R.N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. *American Journal of Physics*, 65, 45-54.
- Redish, E.F., & Wilson, J.M. (1993). Student programming in the introductory physics course: M.U.P.P.E.T. *American Journal of Physics*, 61(3), 222-232.
- Reid, W.M., & Arseneau, D.F. (1971). Labs of unlimited scope. *American Journal of Physics*, 39, 271-274.
- Reif, F., & St. John, M. (1979). Teaching physicists' thinking skills in the laboratory. *American Journal of Physics*, 47(11), 950-957.
- Rosenquist, M.L., & McDermott, L.C. (1987). A conceptual approach to teaching kinematics. *American Journal of Physics*, 55(5), 407-415.
- Shaffer, P.S., & McDermott, L.C. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of instructional strategies. *American Journal of Physics*, 60(11), 1003-1013.
- Shaffer, P.S., & McDermott, L.C. (2005). A research-based approach to improving student understanding of the vector nature of kinematical concepts. *American Journal of Physics*, 73(10), 921-931.
- Smith, T.I., & Wittmann, M.C. (2007). Comparing three methods for teaching Newton's third law. *Physical Review Special Topics – Physics Education Research*, 3(020105).
- Sokoloff, D.R., & Thornton, R.K. (1997). Using interactive lecture demonstrations to create an active learning environment. In E.F. Redish & J.S. Rigden (Eds.), *The changing role of Physics departments in modern universities: Proceedings of ICUPE* (pp. 1061-1074). Melville, NY: American Institute of Physics.
- Sokoloff, D.R., & Thornton, R.K. (2004). *Interactive lecture demonstrations, active learning in introductory physics*. Hoboken, NJ: John Wiley & Sons, Inc.

- Sokoloff, D.R., Thornton, R.K., & Laws, P.W. (2004a). *RealTime physics active learning laboratories module 1: Mechanics* (2nd Ed.) Hoboken, NJ: John Wiley & Sons, Inc.
- Sokoloff, D.R., Thornton, R.K., & Laws, P.W. (2004b). *RealTime physics active learning laboratories module 2: Heat and thermodynamics* (2nd Ed.) Hoboken, NJ: John Wiley & Sons, Inc.
- Sokoloff, D.R., Thornton, R.K., & Laws, P.W. (2004c). *RealTime physics active learning laboratories module 3: Electric circuits* (2nd Ed.) Hoboken, NJ: John Wiley & Sons, Inc.
- Sokoloff, D.R., Thornton, R.K., & Laws, P.W. (2004d). *RealTime physics active learning laboratories module 4: Light and optics* (2nd Ed.) Hoboken, NJ: John Wiley & Sons, Inc.
- Sorensen, C.M., Churukian, A.D., Maleki, S., & Zollman, D.A. (2006). The New Studio format for instruction of introductory physics. *American Journal of Physics*, 74(12), 1077-1082.
- Stelzer, T., Brookes, D.T., Gladding, G., & Mestre, J.P. (2010). Impact of multimedia learning modules on an introductory course on electricity and magnetism. *American Journal of Physics*, 78(7), 755-759.
- Stelzer, T., Gladding, G., Mestre, J.P., & Brookes, D.T. (2009). Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content. *American Journal of Physics*, 77(2), 184-190.
- Steinberg, R.N., Wittmann, M.C., & Redish, E.F. (1997). Mathematical tutorials in introductory physics. In E.F. Redish & J.S. Rigden (Eds.), *AIP Conference Proceedings Vol. 399: The changing role of physics departments in modern universities, proceedings of the International Conference on Undergraduate Physics Education* (pp. 1075-1092). Melville, NY: American Institute of Physics.
- Sweller, J. (forthcoming). Cognitive load theory. In J.P. Mestre and B.H. Ross (Eds.), *The Psychology of Learning and Motivation: Cognition in Education*. San Diego, CA: Elsevier / Academic Press.
- Thacker, B., Kim, E., Trefz, K., & Lea, S.M. (1994). Comparing problem solving performance of physics students in inquiry-based and traditional introductory physics courses. *American Journal of Physics*, 62(7), 627-633.
- Thornton, R. (1989). Tools for scientific thinking: Learning physics with real-time laboratory measurement tools. In E. Redish and J. Risley (Eds.), *Computers in physics instruction* (pp. 177-189). Raleigh, NC: Addison-Wesley.
- Thornton, R.K., & Sokoloff, D.R. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics*, 58(9), 858-867.

Thornton, R.K., & Sokoloff, D.R. (1997). RealTime Physics: Active learning laboratory. In E.F. Redish & J.S. Rigden (Eds.), *AIP Conference Proceedings Vol. 399: The changing role of physics departments in modern universities, proceedings of the International Conference on Undergraduate Physics Education* (pp. 1101-1118). Melville, NY: American Institute of Physics.

Trowbridge, D.E., & McDermott, L.C. (1980). Investigation of student understanding of the concept of velocity in one dimension. *American Journal of Physics*, 48(12), 1020-1028.

Trowbridge, L.W., Bybee, R.W., & Powell, J.C. (2000). *Teaching secondary school science: Strategies for developing scientific literacy* (7th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Trowbridge, D.E., & McDermott, L.C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 49(3), 242-253.

Van Heuvelen, A., & Etkina, E. (2006). *The physics active learning guide*. San Francisco, CA: Addison Wesley.

Vokos, S., Shaffer, P.S., Ambrose, B.S., & McDermott, L.C. (2000). Student understanding of the wave nature of matter: Diffraction and interference of particles. *American Journal of Physics*, 68(S1), S42-S51.

Von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, 80(1), 121-140.

Wilson, J. (1994). The CUPLE physics studio. *The Physics Teacher*, 32, 518-523.

Wittmann, M.C., Steinberg, R.N., Redish, E.F., and the University of Maryland Physics Education Research Group (2004a). *Activity-based tutorials volume 1: Introductory physics, The physics suite*. John Wiley & Sons, Inc.

Wittmann, M.C., Steinberg, R.N., Redish, E.F., and the University of Maryland Physics Education Research Group (2004b). *Activity-based tutorials volume 2: Modern physics*. John Wiley & Sons, Inc.

Wosilait, K., Heron, P.R.L., Shaffer, P.S., & McDermott, L.C. (1998). Development and assessment of a research-based tutorial on light and shadow. *American Journal of Physics*, 66(10), 906-913.

Wosilait, K., Heron, P.R.L., Shaffer, P.S., & McDermott, L.C. (1999). Addressing student difficulties in applying a wave model to the interference and diffraction of light. *American Journal of Physics*, 67(S1), S5-S15.

Yeo, S., Loss, R., Zadnik, M., Harrison, A., & Treagust, D. (2004). What do students really learn from interactive multimedia? A physics case study. *American Journal of Physics*, 72(10), 1351-1358.

Zollman, D., & Fuller, R. (1994). Teaching and learning physics with interactive video *Physics Today*, 47(4), 41-47.

Zollman, D. (1997). Learning cycle physics. In E.F. Redish and J.S. Rigden (Eds.) *The changing role of Physics departments in modern universities: Proceedings of ICUPE* (pp. 1137-1149). Melville, NY: The American Institute of Physics.

Assessment

The development and validation of multiple-choice concept inventories was (and continues to be) an influential area of research in physics education, particularly the development of mechanics and force concept tests: the Mechanics Diagnostic Test (Halloun & Hestenes, 1985), Mechanics Baseline Test (Hestenes & Wells, 1992), the Force Concept Inventory or FCI (Hestenes, Wells, & Swackhamer, 1992), and the Force and Motion Conceptual Evaluation or FMCE (Thornton & Sokoloff, 1998). In particular, the FCI has experienced widespread use to evaluate the effectiveness of instruction within and across several institutions (Hake, 1998). The bank of available concept inventories has expanded to more than 30 physics inventories, including surveys of attitudes and beliefs (see a list at <http://www.ncsu.edu/per/TestInfo.html>).

The process for developing and validating concept tests typically begins with qualitative studies of student difficulties and incorrect responses (or a review of existing research on common misconceptions), drafting and piloting questions, and a series of statistical procedures to evaluate items and scores (see the handbook by Engelhardt, 2009; and Treagust, 1986). For a comparison across several different inventory development methodologies, see Lindell, Peak, & Foster, 2007. Research on assessment has broadened to include several aspects, such as looking at correlations between inventory scores and other measures of performance, comparing scores across multiple populations (culture and gender), and complex models of student learning beyond pre-post scores. Assessment also refers to alternate formats for measuring student understanding such as rubrics and exams, which will also be discussed in this section.

Research Questions

Development and validation of concept inventories *To what extent do scores on a multiple choice concept test agree with students' written responses to open-ended questions? (a measure of test score validity) How does the context and format of concept test questions impact student responses? How do the conditions under which a concept test is administered affect students' performance on the test?* This area of research includes developing and validating concept inventories, methods for analyzing results from concept inventories (Ding & Beichner, 2009), and issues related to context sensitivity and studies of testing conditions (Ding, Reay, Lee, & Bao, 2008).

Comparing scores across multiple measures *What is the relationship between students' scores on the Force Concept Inventory and performance in a physics course as measured by grades? What is the relationship between students' performance on concept tests and measures of their mathematical skills?* Several studies have compared scores on concept inventories to other measures such as SAT scores (Coletta & Phillips, 2007), scientific reasoning ability (Coletta & Phillips, 2005), and scores on a math diagnostic test (Meltzer, 2002). Other possible measures of comparison include grade point average, course grades, or final exam scores. Some studies also compare tests on one inventory to those on another, such as comparing the FCI to the FMCE (Thornton, Kuhl, Cummings, & Marx, 2009).

Comparing scores across multiple populations (culture and gender) *How do females' average scores on the Force Concept Inventory test compare to males' average scores? How do FCI scores compare across cultures (e.g., American and Chinese students)?* There have also been some recent studies comparing inventory scores across multiple social or cultural groups, such as comparing American and Chinese students (Bao et al., 2009a; 2009b) and comparing males and females (Kost, Pollock & Finkelstein, 2009; Lorenzo, Crouch, & Mazur, 2006; Pollock, Finkelstein, & Kost, 2007). There is also some research on gender differences in the use of online homework (Kortemeyer, 2009).

Course exams and homework. *To what extent does performance on multiple-choice questions agree with students' performance on equivalent open-ended problems? How frequently should homework be assigned and graded, if at all? What format of grader feedback on homework is most effective for promoting student learning?* This area includes research on multiple-choice course exams, particularly their equivalence to performance on open-ended problems (H. Hudson & C. Hudson, 1981; Scott, Stelzer, & Gladding 2006), and homework grading (Bao, Stonebraker, & Sadaghiani, 2008).

Rubrics / process assessment. *To what extent do scores on the rubric agree with other measures of student performance (e.g., traditional grading of a problem solution)? What is the level of agreement (consistency) among multiple people using the rubric to score the same student*

papers? Research questions that pertain to the development of rubrics focus on the validity and reliability of scores on the assessment. Examples of physics rubrics include the scientific abilities rubrics developed for the ISLE curricula (Etkina et al., 2006) and problem solving rubrics (Docktor, 2009). There are also assessments under development for problem solving (Marx & Cummings, 2010).

Complex models of student learning *To what extent do predictions from mathematical models agree with measures of student learning? How are student responses to conceptual questions affected by instruction? What is the pattern of student learning throughout a course?* Some researchers are attempting to create mathematical models to describe student learning (Bao & Redish, 2006; Pritchard, Lee, & Bao, 2008) and others are investigating the peaks and decays of learning that occur “between” pre- and post-testing (Heckler & Sayre, 2010; Sayre & Heckler, 2009).

Theoretical Frameworks

Research on the development and validation of assessments is typically guided by theories of test development from fields such as quantitative measurement in education (see Englhardt, 2009 for an introduction to classical test theory; see also <http://www.teststandards.org/>). It is also guided by evolving definitions of reliability and validity (see chapter 3 of Docktor, 2009 for a review). Historically, validity was defined as determining whether an assessment measures what it claims to measure and relied on statistical measures for each “type” of validity, such as content, construct, predictive, and concurrent validity. The current view of validity has shifted away from the “types” paradigm to a unified concept of validity that includes multiple sources of evidence (American Educational Research Association [AERA], American Psychological Association [APA], & National Council for Measurement in Education [NCME], 1999). The most recent edition of the *Standards for educational and psychological testing* describes validity in the following way:

Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests...The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score

interpretations. It is the interpretations of test scores required by proposed uses that are evaluated, not the test itself. (p. 9)

The 1999 edition of the *Standards* cites five primary sources of evidence for validity: evidence based on test content, response processes, internal structure, relations to other variables (external structure), and consequences of testing. For example, the aspect of *content* considers whether the tasks or items on an assessment are representative and relevant to the domain or topic being tested. The aspect of *response processes* considers whether the processes engaged in by a respondent while taking a test are consistent with the assessment developers' intentions. Internal and external structure aspects consider interrelationships among parts of a test and comparisons with other measures of the same construct.

In addition to issues of validity, the *Standards* also address the importance of reliability and utility for designing tests. Reliability refers to the consistency or stability of scores and score interpretations on an assessment, such as agreement among multiple graders for a free-response question. Utility refers to the potential usefulness of the assessment as perceived by its users, such as researchers and instructors.

The development of assessments is also guided by research literature on test design, such as the procedures outlined by Engelhardt (2009) and Treagust (1986). According to these sources, a general procedure for developing concept tests in science includes the following steps: 1) identify knowledge / define the content area in science you want to test; 2) identify student misconceptions associated with that topic by examining the existing literature, or by student interviews and responses to open-ended questions; 3) develop items and pilot test them in two-tier form – asking for response and written reasoning for the response.

Methodology (Data Collection/Sources and Data Analysis)

Contexts. Basic research to identify common student difficulties is typically done outside the classroom, in an experimental or clinical setting with paid volunteers (refer to the section on *Conceptual Change*). During the development of concept assessments, sample items may also be pilot tested with student volunteers, or could be tested in the context of a course. Statistical analyses of quantitative assessment scores are typically conducted in conjunction with a course where the data collected includes student responses to concept inventory questions, homework, or exams.

Participants. Studies of physics assessments frequently involve undergraduate students in large-scale introductory courses (algebra-based or calculus-based) at a single institution over one or two semesters, sometimes longer. Studies conducted with high school students or disciplinary majors are less common. Very few meta-analyses have been conducted to compare concept inventory scores across multiple institutions. One such analysis is Hake (1998), who examined Force Concept Inventory scores for 62 mechanics courses, comparing interactive-engagement instructional strategies to “traditional” instructional methods.

Data Sources and Analysis The early development stages of a concept inventory typically include interviewing students or reviewing existing research literature to identify misconceptions and/or propose assessment items, and pilot testing these items in both a free-response format and a multiple-choice format. After an assessment has been drafted, the statistical analysis of scores (often to measure the validity and reliability of scores on items) typically involves numerical data and quantitative data analysis. Studies comparing performance on multiple measures or across student populations typically utilize statistical measures, such as correlation coefficients.

Findings

Development and Validation of Concept Inventories

Concept inventories used in physics. As stated previously, to date more than 30 physics concept inventories or attitude surveys have been written (see <http://www.ncsu.edu/per/TestInfo.html>). Many are still under development and undergoing validation procedures, and about half of them have been published. The inventories with accessible references are listed below, in chronological order of their publication date.

Published concept inventories and validation studies:

- MDT: Mechanics Diagnostic Test (Halloun & Hestenes, 1985)
- MBT: Mechanics Baseline Test (Hestenes & Wells, 1992)
- FCI: Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992). In the initial article published about the FCI, the authors identified six conceptual dimensions included on the test (Kinematics, Newton’s three laws, the

Superposition Principle, and Kinds of Force). These came into question when factor analysis could not statistically confirm the validity of those categories, prompting a discussion regarding what the FCI actually measures (Heller & Huffman, 1995; Hestenes & Halloun, 1995; Huffman & Heller, 1995).

- FMCE: Force and Motion Conceptual Evaluation (Thornton & Sokoloff, 1998). An analysis by Ramlo (2008) established validity and reliability estimates for the FMCE.
- TUG-K: Test of Understanding Graphs in Kinematics (Beichner, 1994)
- CSEM: Conceptual Survey in Electricity & Magnetism (Maloney, O’Kuma, Hieggelke, & Van Heuvelen, 2001)
- ECS: Energy Concepts Survey (Singh & Rosengrant, 2003)
- DIRECT: Determining and Interpreting Resistive Electric Circuits concept Test (Engelhardt & Beichner, 2004)
- BEMA: Brief Electricity and Magnetism Assessment (Ding, Chabay, Sherwood, & Beichner, 2006)
- R-FCI: Representational variant of the Force Concept Inventory (Nieminen, Savinainen, & Viiri, 2010)
- QMCS: Quantum Mechanics Conceptual Survey (McKagan, Perkins, & Wieman, 2010)

Other assessments:

- LCTSR: Lawson classroom test of scientific reasoning (Lawson, 1978; reproduced in Coletta & Phillips, 2005). This assessment includes questions on conservation, proportional thinking, identification of variables, probabilistic thinking, and hypothetico-deductive reasoning.
- Statistical analysis of scores. In a meta-analysis of FCI scores, Hake (1998) uses a measure he refers to as “normalized gain” to compare pretest and posttest scores, which is examined in more detail in Bao (2006). The average normalized gain for a group of students is computed by taking the average percent gain (average posttest score % - average pretest score %) divided by the possible gain (100% - average pretest score %). Hake’s analysis concluded that courses utilizing interactive engagement methods produced an average normalized gain of 0.48 ± 0.14 standard deviations on the FCI, whereas courses taught using

“traditional” methods produced an average normalized gain of 0.23 ± 0.04 standard deviations. It has become a common practice in PER to compare the normalized gain values obtained in a study to these values reported by Hake (1998).

Although the normalized gain measure is popular, it has received some criticism. Marx & Cummings (2007) suggest revising the gain measure to a proposed measure of “normalized change”. The main distinction between the two measures is that normalized change revises the formula for situations in which the posttest score is lower than the pretest score, and drops students who score 100% on the pretest (a normalized gain of infinity).

Ding & Beichner (2009) identify five statistical means for analyzing test scores to assess the validity and reliability of those scores: classical test theory, factor analysis, cluster analysis, item response theory, and model analysis. Other publications explore these and other analysis methods in greater detail, including item response theory (Lee, Palazzo, Warnakulasooriya, & Pritchard, 2008; Morris et al., 2006; Wang & Bao, 2010), cluster analysis (Smith & Wittmann, 2008), Rasch model based analysis (Planinic, Ivanjek, & Susac 2010), concentration analysis (Bao & Redish, 2001) and model analysis (Bao & Redish 2006).

- *Influence of testing conditions.* A study by Ding, Reay, Lee, and Bao (2008) outlined several testing conditions that affected pre- and post-instruction scores on the Conceptual Survey in Electricity & Magnetism (CSEM) exam. For example, they found that moving the test administration time by a few days or lectures could have a substantial effect on the CSEM pretest scores, since topics of charge and Coulomb force are often discussed in the first week of the course. This is an important consideration for situations in which concept tests are administered outside of a lecture class, such as during a laboratory or recitation section. They also investigated the effect of several different incentives on posttest scores: giving a few points to take the posttest regardless of performance did not seem to influence overall class average; giving students the option to replace an exam score with their posttest CSEM score resulted in a lower response rate but increased average posttest score (perhaps due to sampling from a higher-achieving group of students), and incorporating CSEM questions into the final exam produced a significant increase in posttest scores (Ding et al., 2008). Henderson (2002) observed only a small increase in scores when the FCI posttest was graded versus when it was ungraded, a difference of about half an item or 2%.

- Web-based assessments. Dancy and Beichner (2006) compared performance on a paper-and-pencil version of the FCI to an animated version on the computer. They found that students were better able to understand the intent of the question when viewing it in animated format and this version was more indicative of their conceptual understanding.

Bonham (2008) studied the reliability, compliance, and security of using web-based assessments in an introductory astronomy course. The study found no statistical differences between administering an astronomy concept inventory and an attitude survey on paper or on the Web. The compliance rate increased with the amount of course credit given for completing web-based assessments and the frequency of e-mail reminders. Most students abided by the request not to print, save, or copy portions of the assessments; a small number of students navigated away from the assessment to use other web-based applications, suggesting they might have used web-based resources. The study was unable to detect the use of nonweb resources for completing assessments (such as textbook, notes, or classmate).

- Question context sensitivity. J. Stewart, Griffin, & G. Stewart (2007) studied the effect of changing questions on the FCI on performance. The authors investigated eight types of question alterations: changing a concrete system to abstract (like changing a person pushing on a box to a force exerted on a block), changing the objects in the question to other realistic objects, removing redundant wrong answers, adding a picture, removing a picture, reordering multiple-choice answers, restructuring the order of a group of questions, and making responses symbolic instead of textual. The strongest positive effects were seen for removing a figure, changing the objects in the question, and removing questions from a group. Although altering a question sometimes vastly affected responses on that particular question, these effects tended to balance each other out and resulted in no overall significant change to the total score.

Steinberg and Sabella (1997) report differences in students' answers when an FCI question is administered in multiple-choice format to when it is an open-ended question on an exam. Students typically scored higher on the exam questions, which were less "colloquial" than the FCI and often prompted them to draw the forces acting on objects. Although they observed a correlation between student performance on the two formats, there were several instances in which students gave inconsistent answers. The researchers caution educators to be wary that student understanding is more complex than simple correct or

incorrect answers to multiple choice questions, since changing the context of the test can alter responses. Dufresne, Leonard, & Gerace (2002) also reported specific examples of how changing FCI questions alters students' answers. They describe one example in detail: the situation of two balls of unequal weight dropped from the same height. Although the majority of students surveyed responded correctly to the FCI question that the balls would reach the ground at about the same time, several of these students incorrectly believed that the forces acting on the balls were also equal. The possible influence of context familiarity was briefly addressed by Huffman & Heller (1995), who suggested that the degree of familiarity with particular objects, such as hockey pucks or rockets, might affect a student's scores on those items.

Comparing Scores Across Multiple Measures

- Comparing inventory scores to other measures. The widespread availability of physics concept inventories has prompted some researchers to question whether these tests can be used as placement exams. In other words, is a student's performance on a test like the FCI a predictor of their performance in an introductory physics course? Are strong math skills essential for learning physics? Several studies have compared scores on physics concept inventories to other performance measures such as SAT scores, grade point average, course grades, exam scores, math tests, problem solving, and Lawson's classroom test of scientific reasoning. Meltzer (2002) reported significant correlations between normalized gain on a Conceptual Survey of Electricity (CSE) with pretest scores on an algebra/trigonometry math skills test for three out of four groups ($r=0.38$, 0.46 , and 0.30). Coletta, Phillips, and Steinert (2007) reported strong correlations between normalized gain on the FCI and SAT scores ($r=0.57$ at a high school and $r=0.46$ at a university), and correlations between normalized gain on the FCI and scores on Lawson's classroom test of scientific reasoning ($r=0.51$). Another study claims that math ability is a significant predictor of performance in a physics course (Champagne & Klopfer, 1982; Champagne, Klopfer, & Anderson, 1980). Henderson (2002) emphasizes that the FCI cannot be used as a placement test, because although high-achieving students (scoring higher than 60% on the pre-test) usually get an A or B in the course, the test does not do a good job of predicting failure.

- *Comparing scores on multiple physics inventories.* One study has compared scores on the FCI to scores on the FMCE, which both cover one-dimensional kinematics and Newton's Laws, but measure slightly different content and differ in format (Thornton, Kuhl, Cummings, & Marx, 2009). For example, the FMCE includes questions in pictorial and graphical format whereas the FCI is primarily in verbal and pictorial format, and the FCI is broader in that it includes some questions on two-dimensional motion with constant acceleration. The researchers reported that scores on the FCI were significantly higher than on the FMCE, and there was a strong correlation between scores on the two exams, even though some students showed inconsistent performance on the two tests. They also observed differences in gains on the FCI and FCME for different instructional strategies. For example, *Cooperative Group Problem Solving* showed equal gains on both tests, whereas using *Interactive Lecture Demonstrations* showed higher gains on the FMCE than on the FCI. There is also some inconsistency in observing correlations between diagnostic pre-test scores and normalized gain. Hake (1998) reported no correlation between FCI pre-test scores and normalized gain, and Meltzer (2002) reported no correlation between CSE Pre-test scores and normalized gain, whereas Coletta et al. (2007) did observe a correlation between FCI pre-test scores and normalized gain.

Comparing Scores Across Multiple Populations

- *Cross-cultural comparisons of concept inventory scores.* A comparison of concept inventory scores for Chinese and American students found that Chinese students scored higher on content knowledge tests like the FCI and BEMA, but scored at the same level as Americans on the Lawson Classroom Test of Scientific Reasoning (Bao et al. 2009a; 2009b).
- *Gender comparisons of concept inventory scores.* In introductory calculus-based physics courses for scientists and engineers at large-enrollment institutions, females are underrepresented and typically make up 20-25% of the class (Docktor & Heller, 2008; Kost, Pollock, & Finkelstein, 2009). Recent analyses of test scores in these courses indicate men significantly outperform women on some physics concept inventories, such as the FCI and FMCE, by an average of 10-15%. There are conflicting views from different institutions regarding the extent to which interactive engagement instructional strategies can reduce and/or eliminate the gender gap in scores. Lorenzo, Crouch, & Mazur (2006) claim that interactive engagement instructional strategies such as *Peer Instruction* at Harvard reduce or

eliminate the pre-instruction gender gap in FCI scores. In partially interactive classes the gender gap decreased from 13% before instruction to 7.8% after instruction, and in fully interactive teaching approaches it was essentially eliminated, decreasing from 9.2% to 2.4% (Lorenzo et al., 2006). In contrast, the University of Minnesota observed a persistence of the FCI gender gap on both pre-test and post-test for *Cooperative Group Problem Solving* instruction (Docktor & Heller, 2008). The data examined included 40 fall term classes with more than 5500 students over ten years, showing a gap of $15.3 \pm 0.5\%$ that decreased slightly to $13.4 \pm 0.6\%$ although there was essentially no gender difference in course performance as determined by course grade ($1.5 \pm 0.2\%$). The University of Colorado at Boulder observed a persistence in the gender gap on FMCE scores from pre to post but there were variations by instructor to the extent curricula were implemented, such as ConcepTests, online homework, and *Tutorials in Introductory Physics* (Pollock, Finkelstein, & Kost, 2007). In partially interactive classes the gender gap averaged 10.9% pre-instruction and 13.6% after instruction; in fully interactive, it decreased slightly from an estimated 11.2% to 10.0%. In a second-semester course on electricity and magnetism, there was a small gender gap on the BEMA pre-test ($1.8 \pm 0.5\%$) which increased to $6.1 \pm 1.0\%$ on the post-test. There were some differences in course performance, where females outscored males on homework and participation, whereas males scored higher on exams, resulting in overall similar course grades (Kost, Pollock, & Finkelstein, 2009; Pollock et al., 2007). Kost et al. (2009) used logistic regression analysis to determine that the odds of scoring above 60% on the post-test were not different for males and females when controlling for pre-test scores (FMCE, math, and attitude), suggesting that preparation and background are responsible for the persistence of the gap, not gender explicitly.

Some researchers have attempted to account for pre-test gender differences by examining the preparation of males and females entering introductory college physics courses after high school. Kost, Pollock, & Finkelstein (2009) found that although there is a correlation between years of high school physics and concept inventory post-test scores, gender differences could not be explained by whether students took high school physics because the gender gap was present for matched groups. Hazari, Tai, & Sadler (2007) also found that background factors such as math SAT scores, calculus, and high school grades were equally predictive of success in university physics for both males and females. Factors that

differentially affected male and female performance included high school pedagogy, assessment characteristics, and affective factors such as family attitude toward science and father's encouragement (Harazi et al., 2007; Sadler & Tai, 2001; Tai & Sadler, 2001). Kost-Smith, Pollock, & Finkelstein (2010) suggest that the under-representation of women in physics careers may be due to "an accumulation of small gender differences over time" in aspects such as retention, performance, and attitudes and beliefs such as personal interest (p. 1). A recent intervention involving a brief written exercise on the topic of "values affirmation" successfully reduced the gender gap on FMCE scores at the University of Colorado, suggesting that psychological interventions hold promise for addressing performance differences in science (Miyake et al., 2010).

- *Gender comparisons on homework*. Kortemeyer et al. (2008) determined that female students benefited more from the online homework system CAPA than male students as measured by final course grade. Further study of gender differences in use of the CAPA homework system found that more females than males reported that they consulted peers online or discussed problems with teaching assistants prior to attempting a problem (21% females compared to 9% males) whereas a higher fraction of men reported immediately attempting a problem after reading it (58% males compared to 39% females) (Kortemeyer, 2009). There were no significant differences with respect to the time students reported spending on homework (females averaged 5.6 hours per week, males 5.1 hours per week). Males and females differed in the nature of their comments regarding multiple attempts to enter a solution. Male students stated that when they got a problem incorrect, they would use subsequent attempts to enter "random stuff" or manipulations of an equation, whereas Female students stated that additional attempts gave them an opportunity to explore multiple approaches to the problem without stressing out about grades.

Course Exams and Homework

- *Multiple-choice course exams*. Scott, Stelzer, & Gladding (2009) examined the reliability and validity of using multiple-choice midterm exams in an introductory calculus-based electricity & magnetism course. They compared students' multiple-choice exam scores to written explanations graded by instructors. The multiple-choice scores were found to be highly consistent with instructors' ratings of the students' explanations, indicating that

multiple-choice exams “gave a statistically equivalent assessment of their understanding” (p. 5). Although multiple-choice exams are easier to grade than free-response questions or problems, the authors caution that high-quality multiple-choice questions are more difficult to create. H. Hudson & C. Hudson (1981) also found a high correlation between performance on multiple-choice test questions and free response questions, suggesting that multiple-choice tests can “provide essentially the same information as would hand-graded problems” (p. 838).

- *Homework.* A common practice in introductory physics courses is to assign open-ended problems as homework; however, there are many different approaches for assigning and grading homework (if it is graded at all). For some courses, especially large enrollment courses, there is a concern that grading homework is time-consuming and does not provide timely feedback to students. Bao, Stonebraker, & Sadaghiani (2008) conducted research on a more flexible homework method in which students have some freedom to choose which problems to submit, and half of the problem solutions are available before the homework is due. Bao et al. (2008) found that the majority of students preferred the new approach and believed that early access to the solutions improved their learning. In addition, students who used the new homework method scored significantly higher (5-7%) on the Mechanics Baseline Test than students who used traditional homework methods.

One way to provide students with immediate feedback on their performance might be through the use of electronic scoring via web-based homework systems. As stated in the section on Problem Solving, there are several web-based homework systems in physics including the “computer-assisted personalized assignment” system CAPA (Kashy et al., 1993; Kashy et al. 1995; Kortmeyer et al., 2008; Pascarella, 2002), Mastering Physics, and CyberTutor (Lee, Palazzo, Warnakulassoriya, & Pritchard, 2008). Studies comparing paper-and-pencil homework to web-based homework have found mixed results, with some studies citing no differences in performance (Bonham, Deardorff, & Beichner, 2003) and others citing improved performance with web-based tutors or homework systems (Cheng, K.K., Thacker, B.A., Cardenas, R.L., & Crouch, C., 2004; Dufresne, Mestre, Hart, & Rath, 2002; Morote & Pritchard, 2009; VanLehn et al., 2005).

Lee, Palazzo, Warnakulasooriya, & Pritchard (2008) used item response theory to analyze MIT students’ performance on multiple attempts to answer a question on the

Mastering Physics homework system. Among students who answered an item incorrectly and received tutoring help, 59% were able to get the solution on the second attempt, and another 2/3 of students solved it on their third attempt. Weaker students benefited more from the tutoring – students who consulted hints before submitting an answer (even though there was a penalty for this) showed a large increase of skill, 1.9 standard deviations better. This study also indicated that the hints provided by Mastering Physics improved performance on subsequent attempts of a problem, whereas simply indicating that an attempt was incorrect was insufficient to improve subsequent performance. In a different study, the same researchers found a negative correlation between incidence of homework copying and exam performance (Palazzo, Lee, Warnakulasooriya, & Pritchard, 2010). Students who repetitively copied homework performed worse on exam problems requiring analytic responses. However, there was no significant difference on normalized gain for the Mechanics Baseline Test concept inventory.

Rubrics for Process Assessment

A *rubric* is an assessment tool that indicates a score and the criteria that must be met to achieve that score. It can be comprised of several sub-categories (an analytic rubric) or can be holistic, and is often formatted as a table or grid. Rubrics are used in education research to provide criteria for objective scoring of student responses, but they are also used in classrooms to communicate expectations to students and help them self-assess their own work.

- *Assessing scientific abilities.* Etkina et al. (2006) developed a set of tasks and rubrics to provide students with formative feedback on science-process skills, such as the ability to represent physical processes in multiple ways, design and conduct experimental investigations, and communicate their results. In classes that used the rubrics, students improved their abilities to design an experiment, devise a mathematical procedure to solve an experimental problem, and communicate procedures. In addition, these students performed better than students taught by traditional methods on multiple measures including use of diagrams in problem solving (60% compared to 20%) and performance on multiple choice problems on final tests.
- *Assessing problem solving.* Most measures of problem solving performance given in the classroom focus on the correctness of the end result or partial results, often comparing a

student's written solution to features of an instructor solution. Scoring criteria and point values are often determined on a problem-by-problem basis. Docktor (2009) developed a more descriptive problem solving measure in the form of a general rubric to evaluate students' written solutions to physics problems along five processes: summarizing problem information and generating a useful description, selecting physics principles (physics approach), applying physics principles to the specific conditions of the problem, executing mathematical procedures, and the overall communication or logical progression of the solution. These skill categories were found to be consistent with research literature on problem solving, agree with instructors' beliefs, and are representative of the processes students engage in while solving physics problems (as measured by problem solving interviews). The rubric is applicable to a range of problem types, physics topics, and solution formats.

Another effort to assess problem solving includes recent research by Marx & Cummings (2010) and Cummings & Marx (2010) to develop an assessment of students' ability to solve standard textbook-style problems. The draft instrument includes topics of Newton's Laws, energy, and momentum, with 15 items (five from each topic).

Stewart & Ballard (2010) assessed students' written presentation on open-response questions on four hourly exams. The problems were scored along three categories: math (symbols, relations, and numbers), language (sentences and words), and graphics (i.e., drawings). These presentation elements could explain 32-26% of variance in test performance, but could not explain scores on a conceptual posttest. The incidence of language (sentences and words) and fewer numbers on a student's test was correlated with better performance.

Complex Models of Student Learning

This section includes efforts to extend score-based analyses beyond pretest-posttest score comparisons.

- *Mathematical learning models that consider prior knowledge.* Pritchard, Lee, & Bao (2008) present four mathematical models to describe student learning that differ in their assumptions about the influence of students' prior knowledge (as measured by pre-test scores). The *tabula rasa* model assumes that change in performance (or test gain) is

unaffected by prior knowledge, whereas in the *constructivist* and *connectedness* models assume prior knowledge affects learning, and the *tutoring* model assumes efficient knowledge acquisition and therefore a high rate of learning. The researchers found that data from MIT had a lower value on the “connectedness” parameter than data from the University of Minnesota, indicating that MIT students were using memorization to answer questions on the FCI.

- *Model analysis*. Bao & Redish (2006) propose a method for analyzing concept inventory scores to assess the level of confusion in a class. An individual student’s responses are characterized mathematically as a vector, representing the probability they will apply a particular conceptual model. In addition to indicating how many students answered a question correctly, this model analysis approach indicates a “level of confusion” or the state of a class’s knowledge.
- *Learning progressions*. A few researchers have investigated the progression of learning throughout an introductory electricity and magnetism course, looking at the “peaks” of high scores and “decays” or decreases in performance on conceptual questions that occur between the administration of pre- and post-tests (Heckler & Sayre, 2010; Sayre & Heckler, 2009). They observed a rapid increase in performance that occurred after homework assignments (not after lectures or laboratories) but decreased a few days later. The researchers also observed interference effects, where learning about electrical potential impeded students’ performance on questions related to the vector nature of electric fields, but performance recovered later when learning about magnetic fields.

Strengths & Limitations of Assessment Research

Strengths.

- The topic of assessment using multiple-choice inventories lends itself to large-scale, quantitative studies. Some of the research studies or meta-analyses cited in this section include test scores for upwards of hundreds or even thousands of students, which provide confidence in the interpretation of results.
- The widespread availability of concept inventories in physics has facilitated their use in classrooms both nationally and internationally. Instructors have several tools they can use

to evaluate the effectiveness of their instruction, and can compare published results from other institutions to the scores of their students.

- From the perspective of making physics instruction more effective, concept inventories historically have served a pivotal role in catalyzing dialogs among physics professors who were initially very skeptical that students (especially the ones they taught) could hold erroneous conceptual notions on fundamental physics ideas following instruction.

Limitations

- Many studies that report data on inventories do not give sufficient information about when they were administered and incentive conditions, which can have a significant impact on results (see Ding et al., 2008).
- It is typical for studies to report average scores for a particular class or group of students, but many do not indicate the distribution of scores (e.g., histograms) or separate the results for underrepresented minorities.
- Some research on conceptual tests, especially the Force Concept Inventory, have produced contradictory results. Although the test authors claim the questions and choices were developed based on students' open-ended responses to questions, subsequent studies have found that changing the format of test questions or the context of the questions can dramatically alter students' responses (Dufresne et al., 2002; Steinberg & Sabella, 1997; Stewart et al., 2007).
- Although several sophisticated statistical methods exist for analyzing test scores and have even been described in the PER literature (see Ding & Beichner, 2009 for a review), these procedures have not become widespread. It is possible that the complexity of the statistical methods makes them inaccessible to physics instructors who lack training in PER.

Areas for Future Study

- Homework and exams. There is very little research on homework, such as how much to assign, when to collect it, and how to grade it (if at all). In addition, studies comparing paper and web-based homework have conflicting results, with some stating that web-based homework is better than paper, and some stating that they result in equal learning. There is

also a need for research on the type and quality of feedback provided to students, such as what kind of tutoring is most helpful in web-based homework systems. Similarly, there is a need for research on exams, including the optimal frequency of administering quizzes or tests, what types of questions are best (short answer, essay, multiple choice, problem solving, etc.), and whether comprehensive final exams have merit. Educational psychology research on the “testing effect” suggests that more frequent testing improves learning and retention, but this has not been studied in science classes (McDaniel, Roediger III, & McDermott, 2007). In addition, there exists little or no research on how students prepare for physics exams, or the administration and use of practice exams.

- *Development of additional assessments.* As indicated by the list at: <http://www.ncsu.edu/per/TestInfo.html>, there are several concept inventories under development and only about half of inventories in physics are published. However, there is some need for alternate forms of assessment (other than concept inventories), such as tests for problem solving and mathematics, or measures of expertise such as categorization tasks. A multiple-choice problem solving assessment for text-book style problems is in development (Cummings & Marx, 2010; Marx & Cummings, 2010). The PER community also lacks a consistent math skills exam; several institutions use their own, self-made exam to assess students’ skills. Once such measures exist, additional research could investigate the relationship between scores on concept tests and problem solving performance or the relationship between conceptual knowledge, mathematical skills, and problem solving.
- *Gender and other population differences.* In general, there is a need for studies in PER to report data that is separated for particular populations, such as underrepresented minorities. Research reported in this section indicates that females score lower on average than males on some concept inventories like the FCI and FMCE by as much as 10-15%. The reasons for this gender gap in performance is not well understood and needs further investigation. For example, does the gender gap also exist for electricity and magnetism tests (CSEM or BEMA)? Does it exist for all introductory course populations (non-science, premedicine, engineers, majors, etc.)? How are these results related to the concept of *stereotype threat*? (Miyake et al., 2010)
- *Testing conditions.* There is a need for additional research on testing conditions, such as the influence of when a concept inventory is administered, how it is administered, and

incentives provided to students. Ding et al. (2008) showed for the CSEM test that even a few days or lectures can influence pre-test scores and it would be useful to have similar information for other tests. There have been very few studies mapping the progression of learning, decay, and interference that occur during a course and “between” pre- and post-tests (Heckler & Sayre, 2010; Sayre & Heckler, 2009). Also, the issues involved in administering concept inventories electronically has only been minimally explored (Bonham, 2008).

- *Consistent procedures for analyzing and reporting scores* The “normalized gain” is a commonly reported measure for comparing pretest and posttest scores across populations (Hake, 1998), but the statistical origin for this measure is unclear and alternatives have been suggested (such as “normalized change”). It is unclear why normalized gain is still favored, and the PER community should reach an agreement about how to analyze and report scores from concept inventories.

In addition, the PER community would benefit from handbooks that outline procedures for the development and testing of concept inventories, similar to the guidance in Engelhardt (2009). These handbooks should also incorporate the recent interpretations of validity and reliability from the field of quantitative measurement set forth by the AERA, APA, and NCME (1999) standards for educational and psychological testing (<http://www.teststandards.org/>). Many test developers still follow the validity “types” paradigm, which is inconsistent with the *Standards*.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics*, 74(10), 917-922.

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., Liu, Q., Ding, L., Cui, L., Luo, Y., Luo, Y., Wang, Y., Li, L., & Wu, N. (2009a). Learning and scientific reasoning. *Science*, 323, 586-587.

- Bao, L., Fang, K., Cai, T., Wang, J., Yang, L., Cui, L., Han, J., Ding, L., & Luo, Y. (2009b). Learning of content knowledge and development of scientific reasoning ability: A cross culture comparison. *American Journal of Physics*, 77(12), 1118-1123.
- Bao, L., & Redish, E.F. (2001). Concentration analysis: A quantitative assessment of student states. *American Journal of Physics, Physics Education Supplement*, 69(7), S45-S53.
- Bao, L., & Redish, E.F. (2006). Model analysis: Representing and assessing the dynamics of student learning. *Physical Review Special Topics – Physics Education Research*, 2(010103).
- Bao, L., Stonebraker, S.R., & Sadaghiani, H. (2008). A flexible homework method. *American Journal of Physics*, 76(9), 878-881.
- Beichner, R.J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62(8), 750-762.
- Bonham, S.W., Deardorff, D.L., & Beichner, R.J. (2003). Comparison of student performance using web and paper-based homework in college level physics. *Journal of Research in Science Teaching*, 40, 1050-1071.
- Bonham, S. (2008). Reliability, compliance, and security in web-based course assessments. *Physical Review Special Topics – Physics Education Research*, 4(010106).
- Champagne, A.B., & Klopfer, L.E. (1982). A causal model of students' achievement in a college physics course. *Journal of Research in Science Teaching*, 19(4), 299-309.
- Champagne, A.B., Klopfer, L.E., & Anderson, J.H. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics*, 48(12), 1074-1079.
- Cheng, K.K., Thacker, B.A., Cardenas, R.L., & Crouch, C. (2004). Using an online homework system enhances students' learning of physics concepts in an introductory physics course. *American Journal of Physics*, 72(11), 1447-1453.
- Coletta, V.P., & Phillips, J.A. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics*, 73(12), 1172-1182.
- Coletta, V.P., Phillips, J.A., & Steinert, J.J. (2007). Interpreting force concept inventory scores: Normalized gain and SAT scores. *Physical Review Special Topics – Physics Education Research*, 3(010106).
- Cummings, K., & Marx, J.D. (2010). Beta-test data on an assessment of textbook problem solving ability: An argument for right/wrong grading? In C. Singh, M. Sabella, and S. Rebello (Eds.), *AIP conference proceedings Vol. 1289: 2010 Physics Education Research Conference* (pp. 113-116). Melville, NY: American Institute of Physics.

- Dancy, M., & Beichner, R. (2006). Impact of animation on assessment of conceptual understanding in physics. *Physical Review Special Topics – Physics Education Research*, 2(010104).
- Ding, L., Chabay, R., Sherwood, B., & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical Review Special Topics – Physics Education Research*, 2 (010105).
- Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics – Physics Education Research*, 5(020103).
- Ding, L., Reay, N.W., Lee, A., & Bao, L. (2008). Effects of testing conditions on conceptual survey results. *Physical Review Special Topics – Physics Education Research*, 4(010112).
- Docktor, J.L. (2009). *Development and validation of a physics problem-solving assessment rubric*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.
- Docktor, J., & Heller, K. (2008). Gender differences in both force concept inventory and introductory physics performance. In C. Henderson, M. Sabella, & L. Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference* (pp. 15-18). Melville, NY: American Institute of Physics.
- Dufresne, R.J., Leonard, W.J., & Gerace, W.J. (2002). Making sense of students' answers to multiple-choice questions. *The Physics Teacher*, 40, 174-180.
- Engelhardt, P.V. (2009). An introduction to classical test theory as applied to conceptual multiple-choice tests. In C. Henderson and K.A. Harper (Eds.), *Reviews in PER Vol. 2: Getting Started in Physics Education Research*. College Park, MD: American Association of Physics Teachers. Retrieved August 10, 2010 from <http://www.per-central.org/document/ServeFile.cfm?ID=8807>
- Engelhardt, P.V., & Beichner, R.J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, 72(1), 98-115.
- Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D.T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics – Physics Education Research*, 2(020103).
- Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64-74.
- Halloun, I., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 53(11), 1043-1056.

- Hazari, Z., Tai, R.H., & Sadler, P.M. (2007). Gender differences in introductory physics performance: The influence of high school physics preparation and affective factors *Science Education*, 91, 847-876.
- Heckler, A.F., & Sayre, E.C. (2010). What happens between pre- and post-tests: Multiple measurements of student understanding during an introductory physics course *American Journal of Physics*, 78(7), 768-777.
- Heller, P., & Huffman, D. (1995). Interpreting the Force Concept Inventory: A reply to Hestenes and Halloun. *The Physics Teacher*, 33, 503 & 507-511.
- Henderson, C. (2002). Common concerns about the Force Concept Inventory. *The Physics Teacher*, 40, 542-547.
- Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller *The Physics Teacher*, 33, 502 & 504-506.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141-158.
- Hestenes, D., & Wells, M. (1992). A mechanics baseline test *The Physics Teacher*, 30, 159-166.
- Hudson, H.T., & Hudson, C.K. (1981). Suggestions on the construction of multiple-choice tests. *American Journal of Physics*, 49, 838-841.
- Huffman, D., & Heller, P. (1995). What does the Force Concept Inventory actually measure? *The Physics Teacher*, 33, 138-143.
- Kashy, E., Sherrill, B.M., Tsai, Y., Thaler, D., Weinshank, D., Engelmann, M., and Morrissey, D.J. (1993). CAPA, an integrated computer assisted personalized assignment system *American Journal of Physics*, 61(12), 1124-1130.
- Kashy, E., Gaff, S.J., Pawley, N.H., Stretch, W.L., Wolfe, S.L., Morrissey, D.J., & Tsai, Y. (1995). Conceptual questions in computer-assisted assignments. *American Journal of Physics*, 63(11), 1000-1005.
- Kortemeyer, G. (2009). Gender differences in the use of an online homework system in an introductory physics course. *Physical Review Special Topics – Physics Education Research*, 5(010107).
- Kortemeyer, G., Kashy, E., Benenson, W., & Bauer, W. (2008). Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses. *American Journal of Physics*, 76, 438-444.
- Kost, L.E., Pollock, S.J., & Finkelstein, N.D. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics – Physics Education Research*, 5(010101).

Kost-Smith, L.E., Pollock, S.J., & Finkelstein, N.D. (2010). Gender disparities in second-semester college physics: The incremental effects of a “smog of bias”*Physical Review Special Topics – Physics Education Research*, 6(020112).

Lawson, A.E. (1978). Development and validation of the classroom test of formal reasoning.*Journal of Research in Science Teaching*, 15(1), 11-24.

Lee, Y.-J., Palazzo, D.J., Warnakulasooriya, R., & Pritchard, D.E. (2008). Measuring student learning with item response theory.*Physical Review Special Topics – Physics Education Research*, 4(010102).

Lindell, R.S., Peak, E., & Foster, T.M. (2007). Are they all created equal? A comparison of different concept inventory development methodologies. In L. McCullough, L. Hsu, & P. Heron (Eds.), *AIP conference proceedings Vol. 883: 2006 Physics Education Research Conference*(pp. 14-17). Melville, NY: American Institute of Physics.

Lorenzo, M., Crouch, C.H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom.*American Journal of Physics*, 74(2), 118-122.

Maloney, D.P., O’Kuma, T.L., Hieggelke, C.J., & Van Heuvelen, A. (2001). Surveying students’ conceptual knowledge of electricity and magnetism.*American Journal of Physics PER Supplement*, 69(7), S12-S23.

Marshall, J.A., Hagedorn, E.A., & O’Connor, J. (2009). Anatomy of a physics test: Validation of the physics items on the Texas Assessment of Knowledge and Skills*Physical Review Special Topics – Physics Education Research*, 5(010104).

Marx, J.D., & Cummings, K. (2007). Normalized change.*American Journal of Physics*, 75(1), 87-91.

Marx, J.D., & Cummings, K. (2010). Development of a survey instrument to gauge students’ problem-solving ability. In C. Singh, M. Sabella, and S. Rebello (Eds.)*AIP conference proceedings Vol. 1289: 2010 Physics Education Research Conference*(pp. 221-224). Melville, NY: American Institute of Physics.

McDaniel, M.A., Roediger III, H.L., & McDermott, K.B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom.*Psychonomic Bulletin & Review*, 14(2), 200-206.

McKagan, S.B., Perkins, K.K., & Wieman, C.E. (2010). Design and validation of the Quantum Mechanics Conceptual Survey.*Physical Review Special Topics – Physics Education Research*, 6(020121).

Meltzer, D.E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible “hidden variable” in diagnostic pretest scores.*American Journal of Physics*, 70(12), 1259-1268.

- Mestre, J., Hart, D.M., Rath, K.A., & Dufresne, R. (2002). The effect of web-based homework on test performance in large enrollment introductory physics courses. *Journal of Computers in Mathematics and Science Teaching*, 21(3), 229-251. Norfolk, VA: AACE.
- Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., & Ito, T.A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science*, 330, 1234-1237.
- Morote, E.-S., & Pritchard, D.E. (2009). What course elements correlate with improvement on tests in introductory Newtonian mechanics? *American Journal of Physics*, 77(8), 746-753.
- Morris, G.A., Branum-Martin, L., Harshman, N., Baker, S.D., Mazur, E., Dutta, S., Mzoughi, T., & McCauley, V. (2006). Testing the test: Item response curves and test quality *American Journal of Physics*, 74(5), 449-453.
- Nieminen, P., Savinainen, A., & Viiri, J. (2010). Force concept inventory-based multiple-choice test for investigating students' representational consistency *Physical Review Special Topics – Physics Education Research*, 6(020109).
- Palazzo, D.J., Lee, Y.-L., Warnakulasooriya, R., & Pritchard, D.E. (2010). Patterns, correlates, and reduction of homework copying. *Physical Review Special Topics – Physics Education Research*, 6(010104).
- Pascarella, A.M. (2002). *CAPA (Computer-Assisted Personalized Assignments) in a Large University Setting*. Unpublished doctoral dissertation, University of Colorado, Boulder, CO.
- Planinic, M., Ivanjek, L., & Susac, A. (2010). Rasch model based analysis of the Force Concept Inventory. *Physical Review Special Topics – Physics Education Research*, 6 (010103).
- Pollock, S.J., Finkelstein, N.D., & Kost, L.E. (2007). Reducing the gender gap in the physics classroom: How sufficient is interactive engagement? *Physical Review Special Topics – Physics Education Research*, 3(010107).
- Pritchard, D.E., Lee, Y.-J., & Bao, L. (2008). Mathematical learning models that depend on prior knowledge and instructional strategies. *Physical Review Special Topics – Physics Education Research*, 4(010109).
- Ramlo, S. (2008). Validity and reliability of the force and motion conceptual evaluation. *American Journal of Physics*, 76(9), 882-886.
- Redish, E.F., Saul, J.M., & Steinberg, R.N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.
- Sadler, P.M., & Tai, R.H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, 85(2), 111-136.

Sayre, E.C., & Heckler, A.F. (2009). Peaks and decays of student knowledge in an introductory E&M course. *Physical Review Special Topics – Physics Education Research*, 5(013101).

Scott, M., Stelzer, T., & Gladding, G. (2006). Evaluating multiple-choice exams in large introductory physics courses. *Physical Review Special Topics – Physics Education Research*, 2(020102).

Singh, C., & Rosengrant, D. (2003). Multiple-choice test of energy and momentum concepts. *American Journal of Physics*, 71(6), 607-617.

Smith, T.I., & Wittmann, M.C. (2008). Applying a resources framework to analysis of the Force and Motion Conceptual Evaluation. *Physical Review Special Topics – Physics Education Research*, 4(020101).

Steinberg, R.N., & Sabella, M.S. (1997). Performance on multiple-choice diagnostics and complementary exam problems. *The Physics Teacher*, 35, 150-155.

Stewart, J., & Ballard, S. (2010). Effect of written presentation on performance in introductory physics. *Physical Review Special Topics – Physics Education Research*, 6(020120).

Stewart, J., Griffin, H., & Stewart, G. (2007). Context sensitivity in the force concept inventory. *Physical Review Special Topics – Physics Education Research*, 3(010102).

Tai, R.H., & Sadler, P.M. (2001). Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA. *International Journal of Science Education*, 23(10), 1017-1037.

Thornton, R.K., Kuhl, D., Cummings, K., & Marx, J. (2009). Comparing the force and motion conceptual evaluation and the force concept inventory. *Physical Review Special Topics – Physics Education Research*, 5(010105).

Thornton, R.K., & Sokoloff, D.R. (1998). Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics*, 66(4), 338-352.

Treagust, D. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice items. *Research in Science Education*, 16, 199-207.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15(3), 147-204.

Wang, J., & Bao, L. (2010). Analyzing force concept inventory with item response theory. *American Journal of Physics*, 78(10), 1064-1070.

Cognition

Cognitive psychology is a field of psychology focused on studying mental processes such as problem solving, memory, reasoning, learning, attention, perception and language comprehension. It emerged about 60 years ago and in these ensuing decades the field had matured in terms of theories, methodologies, and traditions. Physics Education Research (PER), on the other hand, is not only a younger field of study, but it is also much more applied than cognitive psychology in two regards: 1) PER focuses on physics, whereas the content topic of research in cognitive psychology is usually not important since the focus is on exploring cognitive processes, and 2) quite often, but not always, a primary goal in PER studies is the improvement of teaching and learning in physics—the improvement of psychology instruction is almost never the goal in cognitive psychology research.

Until recently, PER researchers were trained as research physicists in traditional physics areas and then made the transition to education research and learned to do PER by on-the-job experience. Early PER practitioners had little or no formal training in cognition or education, although sometimes were mentored by other physicists interested in, or practicing PER. What this means is that “cognitive studies” in PER, namely studies exploring cognitive processes related to physics learning or problem solving, have not been common. There is, however, a small body of work with interesting findings that, although not directly applicable to improving classroom instruction, carry important implications for physics learning and problem solving.

Research Questions

The research questions investigated under the cognition category generally fall into the following categories:

***Knowledge and Memory.** How is knowledge organized and accessed / activated? In what ways is knowledge activation influenced by framing and context? What is the role of specific examples in knowledge storage and retrieval? Do experts and novices differ in their organization of knowledge and/or memory?*

Attention. *Do experts and novices attend to different aspects of problems?* (Carmichael, Larson, Gire, Loschky, & Rebello, 2010; Chi, Glaser, & Rees, 1982; Feil & Mestre, 2010); *What do novices attend to when learning from worked-out problems?* (Smith, Mestre & Ross, 2010).

Reasoning and Problem Solving *Do experts and novices approach problem solving tasks differently?* (Chi et al, 1982); *How do experts and novices categorize physics problems?* (Chi et al., 1981; Hardiman, Dufresne, & Mestre, 1989); *What is better for developing problem solving skills, working problems on one's own or studying worked out examples?* (Ward & Sweller, 1990)

Learning and Transfer. *What type of knowledge do students transfer and how does the knowledge transferred depend on context?* (Mestre, 2005; Dufresne et al. 2005; Singh, 2005); *What facilitates and impedes transfer of learning across physics-similar but surface-different contexts?* (Hammer, Elby, Scherr & Redish, 2005; Etkina et al., 2010); *What are possible mechanisms for knowledge transfer and how can they be used to analyze/interpret interview data?* (Rebello et al., 2005). It has been known for some time that transfer of learning across contexts is difficult to achieve (see Mestre, 2003 and *Journal of the Learning Sciences*, 15 (4), 2006 for reviews). Recently a number of studies have explored various aspects of transfer in physics contexts.

Theoretical Frameworks

Given that PER is a relatively young field of study, the theoretical frameworks used in cognitive studies are borrowed from cognitive science and made to fit the occasion. Generally theoretical issues in PER studies analyzing cognitive processes can be characterized in the following ways:

- *Expertise and expert performance.* The goal of studies of expertise is to gain insights on effective reasoning and performance. Although it is recognized that expertise research can help us think about how to make the transition from novice to expert more efficient through the design of instructional interventions, turning novices into experts is not necessarily the primary goal of this body of research. The study of expert performance (for reviews, see Bransford, Brown & Cocking, 1999; Ericsson, Charness, Feltovich, & Hoffman, 2006) has revealed various insights about the way that experts store and apply knowledge, some of which will be summarized in the Findings section. Research on

expert performance has allowed the characterization of knowledge in memory for both experts and novices, and how experts and novices use their knowledge to solve or reason about problems. Studies of expertise typically compare the performance of experts and novices on physics tasks such as problem solving or problem categorization.

- Transfer of learning. Transfer studies focus on exploring what knowledge is deployed to reason about a situation and how the particular knowledge deployed depends on context—that is, how and whether or not changes in context of surface attributes of situations that do not change the underlying structure affect what knowledge is brought to bear and/or how it is used in reasoning. (See chapters in Mestre, 2005 for various studies of this type.)
- Metacognition. Metacognition, which refers broadly to thinking about one's own thinking (Flavell, 1979), is thought to impact learning in that being reflective about one's own thinking processes can improve learning and retention due to deeper processing of the information. (Chi, et al., 1989; for a broad overview, see Bransford, Brown & Cocking, 1999).
- Construction and use of categories. Although it may seem obvious, proficient problem solving is largely the ability to categorize new problems into types that the solver knows how to solve (Chi et al 1982; Hardiman, et al., 1989; Ferguson-Hessler & de Jong 1987). There are several theories regarding the construction and use of categories in learning, such as the prototype view, exemplar view, and rule-based views (for a review of concepts and categories, see Medin, 1989 or Medin & Rips, 2005).
- Attention during processing of information or problem solving. When we look at a situation (e.g., a picture of diagram) or read a problem, we need to be selective in what we attend to since short term memory is limited (see Feil & Mestre, 2010 for a brief review and additional references).
- Knowledge organization and memory. As we gain expertise in a subject, that knowledge is organized in memory in ways that promote quick access/retrieval. The knowledge organization of experts is believed to be hierarchical (see Bransford, Brown & Cocking, 1999). Experts' knowledge is bundled with relevant contexts in which the knowledge can be applied and with procedures for applying it. With their extensive experience in problem solving, experts develop *problem-type schemas*, which consist of representations

of problem categories together with appropriate solution procedures (de Jong & Ferguson-Hessler, 1986; Gick & Holyoak, 1983; Reed, 1993).

- Language. The language people use shapes the way they think (Whorf, 1956). The language physicists use to describe and explain physical phenomena is largely metaphorical. The key here is that metaphors used in physics discourse were originally analogies with inherent limitations (Sutton, 1993). As those analogies were used by physicists, they turned into metaphors that students perceive as literal (Sutton, 1993; Brookes & Etkina, 2007). For example – current flows, energy barrier, force acts. In addition, the grammatical structure of physics statements often leads to ontological confusion (Brookes & Etkina, 2007; 2009). For example, a particle *in* a potential well, weight *of* an object, heat transferred *into* a system.
- Learning by analogy and analogical reasoning. Studies of use of analogy in making sense of phenomena and solving problems indicate analogy is commonly used by people—that is, they tend to look for a similar problem that they know how to solve or that is already worked out and map how the solution to the analogous problem can be applied to solve the new problem. For theories of analogy and analogical transfer see Holyoak & Koh (1987) and Gentner (1989). Theories of analogy use for understanding concepts suggest ways in which analogical scaffolding can be a useful tool for guiding students’ use of representations to understand phenomena (Podolefsky & Finkelstein, 2006; 2007a; 2007b).

Methodology, Data Collection/Sources and Data Analysis

Various methodologies have been used in physics cognitive studies that are commonly used in cognitive psychology:

Contexts. Cognitive studies in physics education are difficult to conduct on a large scale, so they typically take place in a small controlled environment outside of the classroom.

Participants. Participants in cognitive research studies are typically students who know some physics, often paid volunteers who have taken one or two introductory courses in physics. The “experts” in expert-novice studies are often physics faculty or graduate teaching assistants who are experienced with teaching introductory courses.

Data Sources and Analysis Research on cognitive processes often requires descriptive, qualitative sources of data such as interviews or self-explanations to infer students' reasoning processes as they engage in a task. Studies on attention might employ sophisticated data collection tools such as eye-tracking devices to record gaze patterns. Studies of problem categorization can utilize a variety of task formats, as described below.

- Verbal reports from clinical interviews One common methodology for studying cognitive processes in physics education research, and many other disciplines as well, is to conduct interviews of subjects engaged in cognitive tasks, the hope being that verbal utterances reflect, at least in part, what the mind is doing (diSessa, 2007; Ericsson & Simon, 1993; Lee, Russ, & Sherin, 2008). There are differing views on what can be ascertained from subjects' verbal reports. One view is that subjects should simply be instructed to "think aloud" as they perform a task since it is believed that doing so does not change the sequence of thoughts that would occur if the subject were to perform the task in silence (Ericsson et al., 2006). Doing more, such as asking subjects why they make particular choices in performing a task (Nisbett & Wilson, 1977) or asking subjects to self-explain while reading and attempting to understand a solution to a problem (Chi, De Leeuw, Chiu & LaVancher, 1994) may change the accuracy of observed performance in studies of expertise (Ericsson & Simon, 1993). Others argue that cognitive clinical interviews are derivative of naturally occurring forms of interaction and are thus ecologically valid (diSessa, 2007), or that the quality of the information obtained is partly dependent on students' perception of the nature of the discourse interaction (Lee et al., 2008). In PER, interview techniques are often used that attempt to scaffold or cue students to previously learned knowledge (see Rebello et al., 2005 for examples and McDermott, 1984; McDermott & Shaffer, 1992); these interview techniques are useful for learning about ways of structuring effective instructional strategies, as in design research (Confrey 2006; Lamberg & Middleton, 2009).
- Self explanations. Asking subjects to explain to themselves out loud the important features of a problem solution as they study it has been used to study the impact on metacognition on learning (Chi, 2000; Chi, Bassok, Lewis, Reimann, & Glaser, 1989).
- Eye tracking. Eye tracking methodologies (tracking eye gazes by capturing/recording position and duration during cognitive tasks) are quite prevalent in cognitive psychology,

and are used to study what people are paying attention to without disrupting reasoning processes. There is agreement that where the eyes look designate what people are paying attention to, so by following eye fixations one can unobtrusively observe what is being attended to during a cognitive task and thereby draw conclusions about what the mind is processing (see Hoffman & Rehder, 2010 for a good review; Rayner, 1998).

- *Change events or flicker technique.* Methodologies are beginning to be used in PER that borrow from those used in visual cognition (Strand & Mestre, 2010). For example, change event methodologies are used to study a phenomenon known as change blindness. Studies of change blindness (e.g., Simons & Levin, 1998) in psychology typically involve an unsuspecting subject engaged in a task (e.g., giving directions to another person, in this case the experimenter, on how to get to a location on a map), and then unbeknown to the subject something is changed (e.g., the experimenter holding a map is changed to another person holding the same map while the subject is momentarily distracted) to see if the subject notices. The flicker technique alternates between two images on a screen shown for 200 ms with a 100 ms black screen separating them; the two images are slightly different and the time to detect the change is measured to study saliency of the thing being changed. These methodologies attempt to ascertain what is considered salient and thus attended to, and what considered non-salient and neglected, or minimally processed, in a situation or task.
- *Problem categorization.* Problem categorization tasks are common in physics cognitive studies. These tasks ask subjects to identify problems that are solved with a similar approach. Three types of categorization tasks have been used: a) Problem sorting, where subjects place into piles problems that are solved similarly (Singh, 2009; Chi, et al., 1981), b) Problem comparison, where subjects are shown either two, or three problems together and asked to identify which are solved similarly (Hardiman, et al., 1989) or are asked to rate how similar problems are on a Likert scale (Mateycik, Rebello, & Jonassen, 2010). In the two-problem task, subjects are usually asked to state whether or not the problems are solved with a similar approach and to state a reason why. In the three problem task, a model problem is used and two comparison problems, and the subject is asked to identify which of the two comparison problems is solved most like the model problem; this task allows experimenting with competing influences on categorization, such as matching a comparison problem to the model problem on either surface attributes, solution principle/procedure, both, or none, and

c) Multiple choice, where a problem is given followed by five multiple choices for the major principle that would be used to solve it and the task is to select the appropriate one (Leonard, et al., 1996).

Findings

Knowledge and Memory

- *Knowledge activation depends on framing and context.* Framing in anthropology and linguistics (Bateson, 1972; Goffman, 1986; MacLachlan & Reid, 1994; Tannen, 1993), and more recently science education (Hammer, Elby, Scherr & Redish, 2005; Redish, 2004) is used to describe how an individual makes sense out of context and decides what is important and what to pay attention to. Recent research has shown that students can recall different knowledge to reason about a physics situation depending on how they frame the task—that is, how students frame a task affects transfer of learning. For example, if framed in terms of a “school” context, then students tend to draw on formal knowledge, whereas if framed in terms of an out-of-school context, more intuitive or experiential knowledge is brought to bear (Hammer et al., 2005); also, nearly identical tasks can be framed in ways that lead to very different physics knowledge being brought to bear (Dufresne, et al., 2005). In short, transfer of *appropriate* knowledge depends highly on framing and context.
- *Knowledge is tied to specific examples.* Since it is difficult for people to think abstractly, a common approach in teaching physics and other sciences is to present the abstract concepts and then follow with a specific worked example to illustrate how the concepts are applied to solve problems. In cognitive science, Ross (1984, 1987) has shown that if following a worked example students are given a subsequent problem to solve that is somewhat similar to the worked example, then the features of the example’s context become a strong cue for students to use the same method or procedure that they learned in the worked example. That is, students’ understanding of the concept becomes bound up in the particulars of the example that is used to illustrate the principle (Bassok & Holyoak, 1989; Reeves & Weisberg, 1994; Ross, 1989). This phenomenon is known as the “specificity effect,” and has been demonstrated in PER (Brookes, Ross, & Mestre, accepted; Mestre, Ross, Brookes, Smith, & Nokes, 2009).

- *Expert-novice differences in knowledge organization and memory.* Experts chunk related knowledge in clusters, with chunks containing concepts and procedures for applying them, as well as contexts/conditions under which concepts and procedures can be applied. When recalled, chunks are effective for reasoning and solving problems since they contain multiple knowledge, procedural, and contextual units. In contrast, novices' knowledge in memory is more amorphous and does not have the efficient chunked organization of experts' knowledge; when reasoning about a physics situation or solving a problem, novices recall units (e.g., equations, similar contexts) individually making the search and thinking processes more hit-or-miss and simultaneously increasing memory load by creating high demands on short term memory (Chi et al., 1982; Eylon & Reif, 1984; Heller & Reif, 1984; Larkin, 1981, 1983; Larkin, McDermott, Simon & Simon, 1980; for a review, see Ch. 2 in Bransford, Brown & Cocking, 1999).

Attention

- *Attending to relevant features in physics diagrams.* Using a change blindness paradigm, Feil and Mestre (2010) investigated whether novices and experts noticed surreptitious changes made to physics diagrams. Two types of diagram changes were used; in one type, the underlying physics did not change, but in the other type, the underlying physics changed significantly. Findings showed that both experts and novices failed to notice a change if the underlying physics remained the same, but that most experts noticed the change if it altered the underlying physics whereas novices did not, suggesting that experts cue salient features of physics. Yet another study by Carmichael, Larson, Gire, Loschky and Rebello (2010) used eye tracking to explore how experts and novices allocate visual attention to physics diagrams in cases where critical information needed to answer a question was contained in the diagram. They found that experts spend more time looking at thematically relevant areas of diagrams than do novices.
- *Attention in problem solving.* Additional eye-tracking studies using physics tasks are beginning to shed light on what students, and experts, attend to while performing cognitive tasks.
 - *Global versus local attention.* One example is a study by Rosengrant, Thomsom & Mzoughi (2009) in which a small number of novices and two experts answer

questions about circuit diagrams while their eye-gazes are being recorded. In contrast to novices, the eye-gaze pattern of experts showed shifting back and forth between their written work and the circuit diagram, suggesting integration of the solution and diagram; experts' gaze patterns also focused on individual resistors that could be combined as well as the path of the current suggesting global processing of multiple facets of the circuit, whereas novices only looked at proximal resistors that could be combined using the series/parallel rules.

- *Students' attention to conceptual text and mathematical steps.* Another study (Smith, Mestre & Ross, 2010) explored what novices look at while studying worked examples in introductory physics. The initial conjecture was that students would focus on mathematical/procedural information in the worked examples and largely neglect textual/conceptual information because instructors' experience is that students seldom discuss the conceptual underpinnings of a problem's solution, and are at a loss of what to say when asked to discuss the big idea used to solve a problem. Somewhat surprisingly, this study found that students spend a considerable portion of time (about 40%) looking at textual/conceptual information, and that the conceptual and mathematical information in the worked examples was being coordinated (i.e., students did not first look at all the mathematical steps and then all the conceptual discussions, but rather integrated the two). However, little conceptual information that students read was retained, suggesting that students may not have been prepared to learn conceptual knowledge and that this ability may lag ability to learn mathematical procedures for solving problems.
- *Expertise allows fluid and efficient processing of problems/diagrams.* Yet another eye tracking study explored what experienced novices in biology, chemistry and physics look at while processing problems from the New York Regent's exams (Tai, Loehr & Bringham, 2006). Findings indicate that the more expertise that the novices had in a domain, the fewer the gaze fixations needed to process information in specific zones (e.g., problem statement zone, graph or image zone, multiple choices zone) and the fewer the saccades (abrupt, rapid eye-gaze movement between different positions on the screen) between zones. This pattern

suggests that increasing expertise allows individuals to extract and represent salient features in memory without having to return to the zone containing them while coordinating information across zones.

- *Visual attention in physics-relevant tasks.* Other eye-tracking studies have touched on issues peripherally related to PER, such as investigations of troubleshooting of malfunctioning circuits (Van Gog, Paas, Van Merriënboer, 2005), of comprehending malfunctioning mechanical devices (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005), of comprehending how mechanical systems work (Hegarty, 1992), and of the relationship between spatial visualization and kinematics problem solving ability (Kozhevnikov, Motes & Hegarty, 2007).

Reasoning and Problem Solving

- *Expert-novice differences in problem solving processes.* We refer the reader to the Problem Solving section (*Problem Solving Approaches* sub-section) for the salient differences between expert and novice problem solving.
- *Problem categorization.* As discussed in the Problem Solving section (*Categorization* sub-section) experts categorize physics problems according to the major concept or principle that can be applied to solve them, whereas novices rely much more on the surface attributes of the problems to categorize them (Chi et al 1981). Competing influences can affect even experts' ability to cue on the major principle in certain types of categorization tasks. Using a three-problem categorization task Hardiman et al. (1989) showed that expertise is nuanced in that experts can sometimes be fooled by surface attributes, and that some novices can consistently rely on principles in making categorization decisions (Hardiman, et al., 1989).
- *Learning from worked examples.* In the Problem Solving section (*Findings, Worked Examples, Example structure and the worked-example effect*) we discussed the worked-example effect, namely research that suggests that studying worked examples is more effective for learning to solve problems than actual practice in solving problems. There is evidence that studying worked examples can develop schemas better than problem solving practice due to the lower levels of cognitive load that go with studying worked examples, which in turn leaves more resources in short term memory to extract and make

sense of solution strategies (Atkinson, Derry, Renkl, & Wortham, 2000; Sweller, Van Merriënboer, and Paas, 1998). These findings have been replicated in other domains, such as statistics (Paas, 1992) and geometry (Paas & Merriënboer, 1994). The worked-examples effect holds promise for designing pedagogical interventions aimed at improving problem solving skills in physics.

Learning

- *Self-explanations while studying worked examples.* Research conducted by Chi et al. (1989) on how students study worked-out example solutions in textbooks to learn topics in mechanics found that successful students explain and justify solution steps to themselves (self-explain) to a greater extent than poor students. The quality of the explanations also differs; good students refer to general principles, concepts, or procedures which they read in an earlier part of the text, and examine how they are being instantiated in the current example (Chi & VanLehn, 1991). This research suggests that generating self explanations can be an effective pedagogical tool to help students process more deeply the conceptual and procedural aspects of problems solving, and hence lead to enhanced learning than could be achieved by simply reading the worked examples.
- *Analogical reasoning.* Recent work suggests that use of analogies during instruction of electromagnetic waves helps students generate inferences, and that students taught with the help of analogies outperformed students taught traditionally (Podolefsky & Finkelstein, 2006, 2007a). Further, blending multiple analogies in instruction generated better student reasoning compared to instruction that did not use blends or that used standard abstract representations to convey the wave concepts (Podolefsky & Finkelstein, 2007b).
- *Language and symbolic forms.* Research on the role of language showed that making its metaphorical nature transparent for the students helps them apply concepts and solve problems (Brookes, 2006). For example, instead of writing forces as W (weight) or T (tension), students benefit when labeling each force with two subscripts to identify two interacting objects – F_{EonO} (force exerted by Earth on Object) or F_{RonO} (force exerted by Rope on Object). Another example is heat. To help students understand that heat is a process of energy transfer and not energy itself, the term heat can be substituted with

“heating.” Some PER-based curriculum materials are using this new language (Van Heuvelen and Etkina, 2006).

Strengths & Limitations of Cognitive Research in PER

The strengths and limitations of this body of research of research include:

Strengths

- PER cognition research builds upon prior research from cognitive science and hence has a corpus of work from which to draw on for methodologies and theoretical frameworks.
- This type of research helps us learn about human cognition in a complex domain (physics) that requires substantial prior knowledge, reasoning and problem solving skills. As such, this type of research can lead to instructional insights and the eventual design of effective instructional innovations.

Limitations

- The findings from PER cognitive research often are not immediately applicable to improving classroom practice.
- Conducting these types of studies most often requires a pool of subjects who have taken least one or several introductory courses since they need to possess some minimal knowledge of physics. This is to be contrasted with most cognitive psychology experiments that use “toy tasks” that can be learned in a one-hour experimental session.
- Classroom-based cognitive studies involving large numbers of subjects tend to be rare because it is much more difficult to control for extraneous variables that might affect outcomes in these settings; hence, the majority of cognitive PER studies are done in carefully controlled laboratory settings. This may mean that results from a lab based study may yield somewhat different results if one attempts the same study in a realistic setting at a larger scale.

Areas for future study

There are many research areas in cognitive science whose potential has not been explored for learning about physics cognition. For example, psycholinguistics is a large and thriving field of study in cognitive science that draws heavily upon eye-tracking methodology, yet studies of

students' understanding of the language of physics is only now beginning to be explored (Brookes & Etkina, 2007; 2009); similar situations exist with areas such as memory and perception research, and with promising methodologies, such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), where virtually no PER that we are aware of exists. Even those subfields (e.g., visual cognition) and methodologies (e.g., eye-tracking) from cognitive science that are beginning to find their way into PER, the number of existing studies is extremely small. As PER matures and broadens its education of Ph.D.s to include more training in cognitive science, we will see an increase in cognitive studies in PER.

References

- Atkinson, R.K., Derry, S.J., Renkl, A. & Wortham, D.W. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70(2), 181–214.
- Bateson, G. (1972). A theory of play and fantasy. In G. Bateson (Ed.), *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology* (pp. 177-193). New York: Ballantine.
- Bassok, M., & Holyoak, K.J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 153-166.
- Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.
- Brookes, D.T. (2006). *The role of language in learning physics*. Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.
- Brookes, D. T. and Etkina, E. (2007). Using conceptual metaphor and functional grammar to explore how language used in physics affects student learning *Physical Review Special Topics - Physics Education Research*, 3(010105).
- Brookes, D. T. and Etkina, E. (2009). Force, ontology, and language. *Physical Review Special Topics Physics Education Research*, 5(010110).
- Brookes, D.T., Ross, B.H., & Mestre, J.P. (accepted). Specificity, transfer, and the development of expertise. Submitted to *Physical Review Special Topics – Physics Education Research*

- Brookes, D.T., Ross, B.H., & Mestre, J. (2008). The specificity effect: An example from refraction. In C. Henderson, M. Sabella, & L. Hsu (Eds.), *AIP Conference Proceedings Vol. 1064: 2008 Physics Education Research Conference*. Melville, NY: American Institute of Physics.
- Carmichael, A., Larson, A., Gire, E., Loschky, L., & Rebello, N.S. (2010). How does visual attention differ between experts and novices on physics problems? In C. Singh, M. Sabella, and S. Rebello (Eds.), *AIP Conference Proceedings Vol. 1289: 2010 Physics Education Research Conference* (pp. 93-96). Melville, NY: American Institute of Physics.
- Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science, Vol. 5* (pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P. & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.
- Chi, M.T.H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.
- Chi, M.T.H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
- Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence vol. 1* (pp. 7-75). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chi, M.T.H., & VanLehn, K. (1991). The content of physics self-explanations *The Journal of the Learning Sciences, 1*(1), 69-105.
- Confrey, J. (2006). The evolution of design studies as a methodology. In R.K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 135-152). Cambridge: Cambridge University Press.
- de Jong, T., & Ferguson-Hessler, M.G.M. (1986). Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology, 78*(4), 279-288.
- diSessa, A.A. (2007). An interactional analysis of clinical interviewing *Cognition & Instruction, 25*(4), 523-565.
- Dufresne, R., Mestre, J.P., Thaden-Koch, T., Gerace, W., & Leonard, W. (2005). Knowledge representation and coordination in the transfer process. In J.P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 155-215). Greenwich, CT: Information Age Publishing.

Ericsson, K.A., Charness, N., Feltovich, P., & Hoffman, R. (Eds.) (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge: Cambridge University Press.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data* (revised edition). Cambridge, MA: Bradford Books / MIT Press.

Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., & Hmelo-Silver, C.E. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. *Journal of the Learning Sciences*, 19(1), 54-98.

Eylon, B.S. & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognition & Instruction*, 1, 5-44.

Feil, A., & Mestre, J. (2010). Change blindness as a means of studying expertise in physics. *Journal of the Learning Sciences*, 19(4), 480-505.

Ferguson-Hessler, M.G.M., & de Jong, T. (1987). On the quality of knowledge in the field of electricity & magnetism. *American Journal of Physics*, 55(6), 492-497.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.

Gentner, D. (1989). The mechanisms of analogical reasoning. In S. Vosniadou and A. Ortony (Eds.) *Similarity and analogical reasoning* (pp. 199-241). Cambridge: Cambridge University Press.

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

Goffman, E. (1986). *Frame analysis: an essay on the organization of experience*. Boston: Northeastern University Press.

Graesser, A.C., Lu, S., Olde, B.A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & cognition*, 33(7), 1235-1247.

Hardiman, P.T., Dufresne, R., & Mestre, J.P. (1989). The relationship between problem categorization and problem solving among experts and novices. *Memory and Cognition*, 17(5), 627-638.

Hammer, D., Elby, A., Scherr, R., & Redish, E.F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89-119). Information Age Publishing.

Hegarty, M. (1992). The mechanics of comprehension and comprehension of mechanics. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 428-443). New York: Springer-Verlag.

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

Hoffman, A.B. & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility *Journal of Experimental Psychology: General, 139*(2), 319-340.

Holyoak, K.J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15*, 332-340.

Kozhevnikov, M., Motes, M.A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science, 31*, 549-579.

Lamberg, T., & Middleton, J.A. (2009) Design research perspectives on transitioning from individual microgenetic interviews to a whole-class teaching experiment *Educational Researcher, 38*(4), 233-245.

Larkin, J. (1981). Cognition of learning physics. *American Journal of Physics, 49*(6), 534-541.

Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner & A.L. Stevens (Eds.), *Mental models*, (pp. 75-98). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science, 208*(4450), 1335-1342.

Lee, V.R., Russ, R.S., & Sherin, B. (2008). A functional taxonomy of discourse moves for conversation management during cognitive clinical interviews about scientific phenomena. In V. Sloutsky, B. Love, & K. McRae (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 1723-1728). Austin, TX.

Leonard, W.J., Dufresne, R.J., & Mestre, J.P. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems. *American Journal of Physics, 64*(12), 1495-1503.

MacLachlan, G.L. & Reid, I. (1994). *Framing and interpretation*. Portland, OR: Melbourne University Press.

Mateycik, F., Rebello, N.S., & Jonassen, D. (2010). Students' rating of problem similarity as a measure of problem solving expertise. *Proceedings of the 2010 Annual Meeting of the National Association for Research in Science Teaching*.

McDermott, L.C. (1984). Research on conceptual understanding in mechanics. *Physics Today*, 37(7), 24-32.

McDermott, L.C., & Shaffer, P.S. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding *American Journal of Physics*, 60(11), 994-1003.

Medin, D.L. (1989). Concepts and conceptual structure. *American Psychologist*, 44(12), 1469-1481.

Medin, D.L., & Rips, L.J. (2005). Concepts and categories: Memory, meaning, and metaphysics. In K.J. Holyoak and R.G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 37-72). New York: Cambridge University Press.

Mestre, J. (2003). *Transfer of learning: Issues and research agenda*. National Science Foundation Report #NSF03-212. Retrieved October 4, 2010 from www.nsf.gov/pubs/2003/nsf03212/nsf03212.pdf

Mestre, J. (2005). Is transfer ubiquitous or rare? New paradigms for studying transfer. In J. Marx, P. Heron, & S. Franklin (Eds.), *AIP Conference Proceedings Vol. 790: 2004 Physics Education Research Conference* (pp. 3-6). Melville, NY: American Institute of Physics.

Mestre, J.P. (Ed.). (2005). *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age Publishing.

Mestre, J.P., Ross, B.H., Brookes, D.T., Smith, A.D., & Nokes, T.J. (2009). How cognitive science can promote conceptual understanding in science classrooms. In I.M. Saleh & M.S. Khine (Eds.), *Fostering scientific habits of mind: Pedagogical knowledge and best practices in science education* (pp. 145-171). Rotterdam, The Netherlands: Sense Publishers.

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429-434.

Paas, F.G.W.C., & Van Merriënboer J.J.G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach *Journal of Educational Psychology*, 86(1), 122-133.

Podolefsky, N.S., & Finkelstein, N.D. (2006). Use of analogy in learning physics: The role of representations. *Physical Review Special Topics – Physics Education Research*, 2(020101).

Podolefsky, N.S., & Finkelstein, N.D. (2007a). Analogical scaffolding and the learning of abstract ideas in physics: An example from electromagnetic waves *Physical Review Special Topics – Physics Education Research*, 3(010109).

Podolefsky, N.S., & Finkelstein, N.D. (2007b). Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies. *Physical Review Special Topics – Physics Education Research*, 3(020104).

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Rebello, N.S., Zollman, D.A., Allbaugh, A.R., Engelhardt, P.V., Gray, K.E., Hrepic, Z., & Itza-Ortiz, S.F. (2005). Dynamic transfer: A perspective from physics education research. In J.P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age Publishing.

Redish, E.F. (2004). A theoretical framework for physics education research: Modeling student thinking. In E. Redish, C. Tarsitani, & M. Vicentini (Eds.), *Proceedings of the Enrico Fermi Summer School, Course CLVI: Italian Physical Society*. Retrieved October 4, 2010 from <http://arxiv.org/ftp/physics/papers/0411/0411149.pdf>

Reed, S.K. (1993). A schema-based theory of transfer. In D.K. Detterman & R.J. Sternberg (Eds.), *Transfer on Trial: Intelligence, Cognition, and Instruction* (pp. 39–67). Norwood, NJ: Ablex.

Reeves, L.M., & Weisberg, R.W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115(3), 381-400.

Rosengrant, D., Thomson, C., & Mzoughi, T. (2009). Comparing experts and novices in solving electrical circuit problems with the help of eye-tracking. In M. Sabella, C. Henderson, & C. Singh, *AIP Conference Proceedings Vol. 1179: 2009 Physics Education Research Conference* (pp. 249-252). Melville, NY: American Institute of Physics.

Ross, B.H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371-416.

Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-639.

Ross, B.H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 456-468.

Simons, D.J., & Levin, D.T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5(4), 644-649. <http://psych.unl.edu/mdodd/Psy498/simonslevin.pdf>

Singh, C. (2005). Transfer of learning in quantum mechanics. In J. Marx, P. Heron, & S. Franklin (Eds.), *AIP Conference Proceedings Vol. 790: 2004 Physics Education Research Conference* (pp. 23-26). Melville, NY: American Institute of Physics.

Singh, C. (2009). Categorization of problems to assess and improve proficiency as teachers and learners. *American Journal of Physics*, 77(1), 73-80.

Smith, A.D., Mestre, J.P., & Ross, B.H. (2010) Eye gaze patterns as students study worked out examples in mechanics. *Physical Review Special Topics – Physics Education Research*, 6(020118).

Strand, N.E., & Mestre, J.P. (2010, July). Detecting changes to physics diagrams. Presentation at the 2010 Summer Meeting of the American Association of Physics Teachers, Portland, OR.

Sutton, C. R. (1993). Figuring out a scientific understanding. *Journal of Research in Science Teaching*, 30(10), 1215-1227.

Sweller, J., Van Merriënboer, J.J.G. & Paas, F.G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.

Tai, R.H., Loehr, J.F., Brigham, F.J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research and Method in Education*, 29(2), 185-208.

Tannen, D. (1993). *Framing in discourse*. New York: Oxford University Press.

Van Gog, T., Paas, F., & Van Merriënboer, J.J.G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement data and concurrent verbal protocol data. *Applied Cognitive Psychology*, 19(2), 205-221.

Van Heuvelen, A., & Etkina, E. (2006). *The physics active learning guide*. San Francisco, CA: Addison Wesley.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7(1), 1-39.

Whorf, B. L. (1956). Science and linguistics. In J.B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. The M.I.T. Press, Cambridge, M.A.

Attitudes and Beliefs about Learning and Teaching

Students have attitudes, beliefs, and expectations about learning physics that can impact the way they behave and perform in a physics course (Halloun, 1997; May & Etkina, 2002; Perkins et al., 2005). For example, a common student belief is that physics is made up of several unrelated pieces of information. As a result, many students approach physics by memorizing formulas without connecting them to a broader understanding of the underlying concepts and principles. This section summarizes frameworks that have been developed to describe students' beliefs and multiple-choice attitude surveys to explore changes that result from instruction.

In addition to students' attitudes and beliefs, the beliefs that instructors have about how students learn can impact their instructional decisions and classroom interactions. Research from PER and science education in general indicates that oftentimes, instructors' beliefs and practices are inconsistent with each other. This section also summarizes research on physics instructors' beliefs, their decisions to adopt research-based curricula, and how PER-based curricula are actually implemented in the classroom.

Research Questions

Key areas of research on this topic include student attitudes and beliefs about learning physics, instructor beliefs about how students learn physics, instructors' decision-making processes, the ways in which instructors implement reformed curricula and instruction, and research on the attitudes and practices of teaching assistants.

Student attitudes and beliefs about learning physics *What attitudes, beliefs, and expectations do students have about learning physics? How do students' attitudes and beliefs change as a result of physics instruction? What instructional strategies are effective for promoting productive attitudes and beliefs?* There exists research on students' beliefs about learning physics (Hammer, 1994; 1995; 1996), and on the impact that their beliefs have on their performance as measured by concept inventory scores and course grades (Halloun, 1997; May & Etkina, 2002; Perkins et al., 2005). Research also exists on instructional strategies designed to explicitly promote productive attitudes and beliefs (Elby, 2001; Hammer & Elby, 2003). Below we review several surveys that

have been developed to measure student attitudes and beliefs before and after instruction, including the MPEX (Maryland Physics Expectation Survey; Redish, Saul, & Steinberg, 1998), VASS (Views About Science Survey, Halloun, 1997; Halloun & Hestenes, 1998), EBAPS (Epistemological Beliefs about Physical Science; Elby, 2001; White et al., 1999), CLASS (Colorado Learning Attitudes about Science Survey; Adams et al., 2006), and AAPS (Attitudes and Approaches to Problem Solving survey; Mason & Singh, 2010).

Faculty beliefs and values about teaching and learning *How do instructors believe students learn physics (e.g., physics problem solving?) What prior knowledge or skills do instructors expect of their students?* (Yerushalmi et al., 2007) *How do instructors view their role as a teacher? What factors contribute to their instructional decisions?* (Henderson & Dancy, 2007). Recent research has explored beliefs and values that faculty have about teaching and learning physics in general (Henderson & Dancy, 2007) and problem solving specifically (Yerushalmi, Henderson, P. Heller, K. Heller, & Kuo, 2007; Henderson, Yerushalmi, Kuo, K. Heller, & P. Heller, 2007; Henderson, Yerushalmi, Kuo, P. Heller, & K. Heller, 2004), including instructors' reasons for choosing problem features (Yerushalmi, Cohen, K. Heller, P. Heller, & Henderson, 2010).

Instructor implementations of reformed curricula *How prevalent is the use of research-based instructional strategies by physics faculty in the United States? How is [research-based strategy X] typically implemented? What is the nature of faculty-student interactions during Peer Instruction? What instructional practices foster an interactive, collaborative classroom environment?* Some very recent studies have explored faculty adoption of PER-based instructional strategies and materials (Dancy & Henderson, 2010; Henderson & Dancy, 2008; 2009), how PER-based strategies like *Peer Instruction* are actually implemented in the classroom (Turpen & Finkelstein, 2009; 2010), and student perceptions of classroom reforms.

Teaching Assistants *What attitudes and beliefs do teaching assistants have about student learning? How do teaching assistants view their role as recitation and/or laboratory instructors? How do teaching assistants impact the success of reformed instruction?* A few studies have explored the attitudes and beliefs that teaching assistants have about student learning, and how those relate to their "buy-in" to reformed instructional strategies and curricula (Goertzen, Scherr,

& Elby, 2009; 2010). Other research explores how instructional style (specifically TA coaching) influences student learning during recitation (Koenig, Endorf, & Braun, 2007).

Theoretical Frameworks

Research studies of attitudes and beliefs about teaching and learning do not always explicitly identify theoretical frameworks. Studies might simply cite previous research findings from education research or psychology as a basis for their work. Some of the PER research on faculty beliefs takes this approach (see for example, Yerushalmi et al., 2007). One exception is research on teaching assistants (Goertzen et al., 2009; 2010) which identifies “Framing” theory from anthropology and linguistics as a basis for understanding how TAs make sense of their teaching experiences. The following examples are frameworks that have been developed and published by researchers in physics education.

Student Attitudes and Beliefs

Hammer (1989; 1994) developed a framework for characterizing students’ approaches to learning physics, which had its basis in education research on learning “orientations” in addition to interviews he conducted with physics students. The three categories of his framework are: student beliefs about the structure of physics (made up of pieces vs. coherent), beliefs about physics content (formulas vs. concepts), and beliefs about learning physics (receiving knowledge vs. actively making sense of new information). Hammer’s framework was later used to design multiple-choice attitude surveys.

Instructional Practices and Conceptions

Dancy and Henderson (2007) introduce a framework for describing instructional practices and conceptions. The framework was developed from research literature, discussions within the PER community, and interviews with physics faculty. The framework’s ten categories of conceptions include: learning view (transmissionist vs. constructivist), beliefs about expertise, knowledge, and the nature of physics; perceptions of their role as an instructor and the role of the school; desired outcomes (goals) for students, diversity, and scientific literacy. The ten categories of practice identify behaviors typical of traditional instruction compared to alternative instruction, and include: interactivity (passive lecturing vs. active student conversation), instructional decisions (made by teacher alone vs. with students), knowledge sources (students receive vs. construct knowledge), learning mode (competitive vs. cooperative), motivation

(external vs. internal), assessments (knowledge-based vs. process-based), content (facts and principles only vs. also teach thinking and problem solving skills), instructional design (knowledge-driven or student-driven), and problem solving (formulaic vs. creative).

Methodology (Data Collection/Sources and Data Analysis)

Contexts. Studies involving students, such as identifying students' attitudes and beliefs about learning physics and the development of attitude surveys, have primarily taken place in conjunction with introductory physics courses. In contrast, studies involving instructors (faculty or teaching assistants) often employ interviews that take place in a controlled experimental setting outside the classroom, or surveys that are administered online.

Participants. Studies of students' attitudes and beliefs about learning physics have typically involved students enrolled in an introductory course. Studies of instructor's beliefs about learning and teaching physics often involve faculty (from a variety of institutional types) who are currently teaching a course or who have substantial experience teaching physics. A very small number of studies involve graduate teaching assistants in their first or second year of graduate school.

Data Sources and Analysis

- Clinical interviews. Research to identify the attitudes and beliefs of students and instructors typically begins with in-depth interviews that are audio and/or video recorded. The transcribed statements are coded by themes that are either determined by the researchers or from existing theoretical frameworks. These methods were used by Hammer (1994; 1995; 1996) to describe students' beliefs about learning physics, and by several researchers investigating instructor's beliefs (Henderson & Dancy, 2007; Henderson et al. 2004; 2007; Yerushalmi, et al. 2007; 2010). Sometimes these instructor interviews include specific *artifacts* such as sample student or instructor solutions that provide a basis for discussion.
- Surveys. Web-based surveys have been used in research on faculty knowledge and use of research-based instructional strategies (Dancy & Henderson, 2010; Henderson & Dancy,

2009). The results are analyzed using descriptive statistical measures, such as reporting the average fraction of survey respondents who reported using a particular research-based instructional strategy.

- *Classroom observations.* Observations of instruction have been used in studies of teaching assistant behaviors (Goertzen, Scherr, & Elby, 2009; 2010). These videotaped observations are analyzed along several different measures, but typically attempt to characterize the nature of student-student and teacher-student interactions.
- *Analyzing students' written work.* In some studies, students' degree of "sophistication" in their beliefs about learning physics is analyzed by scoring their written responses to reflective questions (Kortemeyer, 2007; May & Etkina, 2002). After the researchers have obtained such a quantitative measure, it is analyzed further by statistical correlations with other course measures such as concept inventory scores, course exams, homework, or final course grades.

Findings

Student Attitudes and Beliefs about Learning Physics

This section discusses research on student attitudes and beliefs and the impact these have on learning physics. In these studies, a student's approach to learning is discussed using a variety of terms, such as attitudes, beliefs, values, expectations, views, personal interest, learning orientation, motivation, and epistemological beliefs. Our use of the terms "attitudes and beliefs" is intended to broadly encompass this wide range in terminology. When the term *epistemology* is used, it should be interpreted to refer to "beliefs about what constitutes knowledge in physics and how, as a student, one develops that knowledge" (Hammer, 1994 p. 152). Some researchers have collectively referred to students' attitudes, beliefs, and assumptions about learning physics as their "expectations" upon starting a course (Redish, Saul, & Steinberg, 1998). For a more comprehensive review of research on epistemology and relationship to teaching and learning, see Hofer (2001) and Hofer & Pintrich (1997).

- *Epistemological beliefs in physics.* Hammer (1994; 1995; 1996) developed a framework to characterize students' attitudes and beliefs about physics in terms of three scales: Pieces-Coherence, Formulas-Concepts, and By Authority-Independent (Hammer, 1995; Redish,

2003). The *coherence* measure refers to a student's view of the structure of physics as consisting of several unrelated pieces of information (unfavorable) or as a connected system that is relevant to everyday experiences in the real world (favorable). A student's belief about *concepts* refers to whether they focus on memorizing and blindly applying equations (unfavorable) or attempt to understand the concepts that underlie the equations (favorable). A student with a favorable attitude on the *independence* measure actively attempts to construct her/his own understanding, whereas a student who just receives information from an authority (e.g. instructor or textbook) without question has an unfavorable learning attitude.

Research by Elby (1999) extended Hammer's ideas to explore students' expectations for a physics course and the relationship between epistemological beliefs and study habits. He determined that although most students know what it means to understand physics deeply and how to go about learning physics in a deeper way, they do not engage in productive study habits because they don't see this level of understanding as essential for obtaining good grades in a physics class.

- *Surveys of student attitudes and beliefs.* In order to better characterize student attitudes and beliefs before, during, or after a course, researchers developed an assortment of attitude survey instruments. One of the first surveys developed was the MPEX (Maryland Physics Expectation Survey), a 34-item Likert-scale questionnaire (disagree-agree) that probes six dimensions of students' expectations about learning physics (Redish, Saul, & Steinberg, 1998). The first three dimensions match those identified by Hammer (1994): independence, coherence, and concepts; and there are three additional dimensions: reality link (view physics as relevant to real life), math link (math is a way to represent physical phenomena, not just a tool for calculation), and effort (what students perceive they need to do to be successful in a physics class). Items are scored as being favorable if they agree with how an expert physicist would answer the item, or unfavorable if the response disagrees with expert responses on that item. Upon administering the test to physics students at six universities or schools, the authors consistently observed a decline in students' attitudes following a semester of instruction (Redish et al., 1998). The authors suggest that some physics courses might actually reinforce students' inappropriate approaches to learning by placing an emphasis on memorization rather than a deeper conceptual understanding of the subject matter.

The VASS - Views About Sciences Survey (Halloun, 1997; Halloun & Hestenes, 1998) is another survey instrument that probes students' views of learning science along six dimensions; three scientific dimensions (structure or coherence, methodology or use of mathematics, and validity of scientific knowledge) and three cognitive dimensions (learnability or personal effort, reflective thinking, and personal relevance). The authors use scores on the VASS to classify someone as having expert-like views, mixed (transitional) views, or folk views. For the populations studied, high school and college students scored similarly on the survey, and very few (no more than 10%) were classified as having expert-like views.

The EBAPS survey (Epistemological Beliefs Assessment for Physical Science) is formatted in a slightly different way than other surveys of attitudes and beliefs (<http://www2.physics.umd.edu/~elby/EBAPS/home.htm> Elby, 2001; White et al., 1999). Although part one of the survey includes Likert-style statements (disagree-agree), part two of the survey includes multiple choice questions in which students must decide which statement they agree most with. The five subscales of the EBAPS include: structure of scientific knowledge, nature of knowing and learning, real-life applicability, evolving knowledge, and source of ability to learn (natural ability vs. effort-based). In contrast to the MPEX, the EBAPS survey attempts to only assess epistemological beliefs separately from expectations about learning physics.

The most recently-developed attitude survey is a 42-item questionnaire adapted from previous work, called the CLASS – Colorado Learning Attitudes about Science Survey (Adams et al., 2006). It differs from previous surveys in that the dimensions are constructed from a factor analysis of students' scores rather than a priori expectations of question groupings, and it explicitly considers attitudes toward problem solving. The eight categories include: real world connection, personal interest, sense making / effort, conceptual connections, applied conceptual understanding, problem solving general, problem solving confidence, and problem solving sophistication. There are also six items that do not fit into these categories and are not scored. One study of scores on the CLASS (Gray, Adams, Wieman, & Perkins, 2008) asked students to respond how they think a physicist would answer the questions, and how they personally would answer it. When answering from the perspective of a physicist, students were very accurate at selecting the “favorable” attitude

responses, however they did not answer in the same way when responding with their own personal beliefs.

Gire, Jones, & Price (2009) used the CLASS survey to measure the epistemological development of physics majors. They administered the attitude survey to physics majors who were at different stages in their academic career, including students in their first, second, third, or fourth year of college, and graduate students in physics. They found that among students in their first year of college, the physics majors had more “expert-like” responses on the CLASS than did their introductory physics course peers, who were primarily composed of engineering students. Students in years one, two, and three had similar overall fraction of favorable scores on the CLASS, whereas the scores were slightly higher for students in their fourth year or in graduate school.

A recently published measure assesses shifts in students’ physics course expectations in response to SCALE-UP orientation and instruction (Student Centered Active Learning Environment for Undergraduate Programs). The assessment is called the Pedagogical Expectation Violation Assessment (PEVA) (J. Gaffney, A. Gaffney, & Beichner, 2010). At the beginning of the course, most students expected to attend lectures in an amphitheater classroom, to attend a separate laboratory section, to read the text, and to memorize equations, all with limited opportunities to interact with instructors and peers. As a result of a brief orientation to the course, most students shifted their expectations to be closer to the actual design of SCALE-UP (decreased lecture, an integrated laboratory environment, and more interactions including collaborative group work). The students also reduced their expectation for memorizing at two of three institutions studied.

- Mason and Singh (2010) refined an instrument to measure attitudes and beliefs about problem solving, called the Attitudes and Approaches to Problem Solving (AAPS) survey. They compared responses on the survey across several groups: introductory physics and astronomy students (algebra-based and calculus-based), graduate students, and faculty. The graduate students answered the survey twice, once from the perspective of solving introductory physics problems, and again from the perspective of solving graduate-level physics problems. In general, introductory students had beliefs about problem solving that were less expert-like than the beliefs of graduate students or faculty. When graduate students took the survey from the perspective of solving graduate-level physics problems, their responses were less expert-

like than when they answered the survey for introductory problems. For example, graduate students indicated that they have difficulty checking whether their answers to graduate problems are correct, and they feel they need to seek help from others when they get stuck. *Relationships between epistemological sophistication and performance.* Some studies have explored the relationship between epistemological beliefs and understanding of physics concepts as measured by performance on concept inventories or other course measures. Halloun (1997) reported a significant correlation between scores on the VASS, FCI gain, and course grades. For example, students who were classified as an “expert” profile on the VASS survey were most likely to be high achievers in the physics class.

Perkins et al. (2005) observed a statistically significant correlation between scores on some (but not all) categories of the CLASS survey and pre- and post-test scores on the FMCE (Force and Motion Conceptual Evaluation). The categories of *conceptual understanding* and *math physics connection* were significantly correlated with FMCE (coefficients in the range of 0.20 to 0.30), whereas real world connection, personal interest, and sense making / effort were not significant (correlation coefficients ranging from 0.02-0.17). In addition, when students were grouped by normalized learning gains on the FMCE, students in a high gain bin (>0.9) tended to have more favorable beliefs whereas students with the lowest gain (<0.2) had unfavorable beliefs that declined after the course.

Kortemeyer (2007) investigated statistical relationships between several different measures: epistemology as coded from students’ online discussion behavior associated with web-based homework and scores on the MPEX survey; and physics learning as measured by the FCI, final exam, and course grades. The study found significant correlations between FCI post-test scores, FCI gain, course grades, and the extent to which students’ online discussion posts were physics-oriented (positive) as compared to solution-oriented (negative). These correlations were not observed for the MPEX, suggesting that online discussions might be a useful (and more authentic) diagnostic tool for assessing students’ approaches to learning physics than survey instruments.

May & Etkina (2002) required introductory physics students to submit weekly reports in which they reflected on how they learned particular physics topics in electricity and magnetism. The researchers compared the quality of students’ reflections to their performance on several concept inventories: the Force Concept Inventory (FCI), Mechanics Baseline Test

(MBT), and Conceptual Survey of Electricity & Magnetism (CSEM). They found that students with high pre-post normalized gains had a higher level of epistemological sophistication, such as asking insightful questions and trying to make sense of the material rather than saying they learned formulas.

Lising and Elby (2005) conducted a case study on a single student “Jan” to investigate how epistemology affects learning. They observed that Jan had a separation between formal, classroom reasoning and everyday reasoning which contributed to her difficulties with learning physics. Although she could reason in both ways, she often did not attempt to connect them and/or reconcile inconsistencies. The authors suggest that research-based curricula can be made even more effective by making epistemology an explicit part of instruction, especially through reflective questions on assignments or during class discussions.

- *Instructional strategies to improve students’ epistemological beliefs* Elby (2001) and Hammer and Elby (2003) developed a set of instructional materials and strategies to address students’ epistemological beliefs, which were found to result in significant, favorable shifts on the Maryland Physics Expectation (MPEX) survey. Effective instructional practices included: assigning essay questions in which students must argue for or against multiple perspectives, asking students to reflect on whether their answers agree with their “intuition” during laboratory activities, and submitting journal-like paragraphs in which students reflect about the strategies they use to learn physics (memorization, summarizing the text, solving problems, etc.). The authors acknowledge that implementing these approaches was at the expense of reduced content coverage, they only minimally used the textbook, and they attempted to keep lesson plans flexible to allow time to address students’ difficulties.
- *Gender differences on attitude surveys*. One study of scores on the CLASS (Colorado Learning Attitudes about Science Survey) indicate that females have less expert-like beliefs in statements related to categories of real world connections, personal interest, problem solving confidence, and problem solving sophistication than their male peers (Adams et al., 2006). However, their responses to questions in the sense-making / effort category are slightly more expert-like than those of male students.

Faculty Beliefs and Values about Teaching and Learning

Faculty beliefs and values about how students learn physics influence their decisions about what and how to teach, but due to external factors these beliefs may not be reflected in their actual teaching practices (Henderson & Dancy, 2007; Yerushalmi et al., 2007). An important first step to understanding instructional decisions is to identify a common set of beliefs held by instructors and the situational factors that limit their use of reformed curricula or teaching methods. This information can then be taken into consideration by curriculum designers and leaders of professional development programs when disseminating research-based instructional tools.

- *Faculty conceptions and instructional practices.* Henderson & Dancy (2007) report the results of semistructured interviews with five experienced, tenured physics faculty from four different institutions. They analyzed the faculties' conceptions about teaching and learning and self-reported instructional practices according to a framework outlined in Dancy & Henderson (2007). The ten categories used to rate the conceptions focus on views of student learning and how the instructor views his/her teaching role, which were scored on a scale from being traditional (transmissionist) in nature to more alternative (constructivist). Similarly, the ten categories used to rate self-described instructional practices considers whether an instructor describes his/her actions as being more consistent with alternative instruction (e.g., active, cooperative, creative, with process-based assessments) or traditional instruction (e.g., passive, individualistic, formulaic, with knowledge-based assessments). Henderson & Dancy (2007) found that most of the faculty had views about teaching and learning that were rated as semialternative or a mix between alternative (reformed) views and traditional views. In contrast, their descriptions of their own teaching practices were rated as more traditional in nature. The study concluded that even though these faculty members were familiar with research-based instructional methods, agreed with reform approaches, and had access to materials, they often did not implement them because of external factors beyond their control. These barriers included a need to cover content (and limited time in which to do so), lack of time to prepare for teaching, inappropriate class size and room layout, departmental "norms" for how classes are taught, and student factors (student resistance or poor student attitudes).

- *Beliefs and values about teaching and learning of problem solving.* Problem solving is a primary component to most university physics courses, and a substantial portion of students' grade is often dependent upon their ability to solve problems as homework and on exams. Students are provided with worked-out examples in their textbooks and might observe additional demonstrations of problem solving during class; however, the problem features, example solutions, and problem solving procedures used by different instructors may vary widely. Yerushalmi et al. (2007) and Henderson et al. (2007) used structured interviews with six physics faculty to explore factors that contributed to their instructional decisions related to the teaching and learning of problem solving. The interviews took place with faculty at a research university and utilized specific "artifacts" as a basis for discussion, including different formats for the same problem statement, multiple student solutions, and multiple versions of an instructor solution (Henderson et al., 2007). During the interviews, the participants were prompted to comment on their format preferences for each type of artifact and indicate what they typically use in the classes they teach and why.

The researchers concluded that these physics instructors had unstable, often conflicting beliefs that were constructivist in nature, while their actions in the classroom reflected a traditional model of transmitting information (Yerushalmi et al., 2007). For example, the faculty expressed a belief that students must be reflective learners and solve a lot of problems on their own to gradually build up an understanding of physics (an inquiry-based, constructivist view), but they often provided explicit guidance in course materials such as using problems that are broken into parts (parts a, b, c, etc.) to direct students through a problem solving procedure. In addition, the faculty experienced a conflict between their "physicist" view that places value on compact, concise problem solutions and a "teacher" value of wanting students to communicate their reasoning. As a result, the instructors were unwilling to penalize a student who wrote a very sparse answer that could be interpreted as a correct result, but would penalize students who wrote incorrect reasoning that led to a correct numerical result (Henderson et al., 2004). The faculty who were interviewed acknowledged that their preferences do not match what they use in their courses, often because it takes too much time and expertise to construct high-quality problems and problem solutions, and they also don't want to overwhelm students with complicated or overly detailed problem solutions.

A later study extended this research to include interviews with 30 physics faculty at a wide range of institutions, and focused on instructors' choice of features when writing or selecting physics problems for their class (Yerushalmi et al., 2010). During interviews, faculty indicated their preferences for or against particular problem features, including problems that were qualitative (no calculation), multiple choice, broken into parts, had a real-world context, were wordy (extra unnecessary words), included a drawing, or were complex (required multiple principles for a solution). The interview statements were coded along three dimensions: whether a particular problem feature supported or hindered an instructor's teaching goals, whether the feature was used by a faculty member in their courses in any context, and whether the feature was used on exams. The researchers concluded that for four of these features, the instructors' values were in conflict with their practices. Although faculty believed conceptual questions promote student understanding and a real-world context provides necessary motivation for students, they chose not to use problems with those features in class. In contrast, they stated that problems which are broken into parts and include a drawing hinder their goal that students will learn general problem solving skills, but they frequently use these features to make a problem clear, especially on exams.

Instructor Implementations of Reformed Curricula

This section reviews findings from studies related to faculty adoption of PER-based instructional strategies and materials, how PER-based strategies are actually implemented in the classroom, and student perceptions of classroom reforms.

- *Self-reported knowledge and use of reformed instruction.* Henderson and Dancy (2009) conducted a large-scale online survey of physics faculty in the United States. Of the 722 faculty who responded, most of them had heard of at least one research-based instructional strategy (87.1%) and about half are familiar with six or more strategies. In terms of use, approximately half of the faculty reported using reformed instruction in their teaching, but often with significant modifications. From the list of 24 research-based instructional strategies, the most common instructional strategy was *Peer Instruction*, with 63.5% of faculty reporting they know about the strategy and 29.2% of them saying they use it in some form. Other strategies that were familiar to more than 40% of faculty

include Physlets, Cooperative Group Problem Solving, Workshop Physics, Just-in-Time Teaching, Tutorials, and Interactive Lecture Demonstrations. Although this survey shows that there is high knowledge of PER-based instruction among faculty, situational factors such as limited time or lack of support deter faculty from using them (Dancy & Henderson, 2010; Henderson & Dancy, 2007). During interviews with five physics faculty, Henderson and Dancy (2008) also found that several faculty expressed dissatisfaction with their interactions with educational researchers. The authors suggest that PER should change its model for disseminating information and materials to better help faculty adapt the strategies to their specific situation.

- *Implementations of Peer Instruction*. Turpen and Finkelstein (2009) used detailed classroom observations to identify variations in the ways that physics instructors implement *Peer Instruction* (Mazur, 1997). They found that the basic elements of *Peer Instruction* were present in all six classes studied: the presentation of conceptual information (usually in the form of questions) and opportunities for students to discuss physics with their classmates. The instructors differed, however, in the average number of conceptual questions asked per class period (range from 3 to 8), how they interacted with students (both during the “voting” time and discussing the responses), and the time spent explaining the solution (range from 1 minute to nearly 4 minutes). One of the most prominent differences was in the average number of explanations heard from students during a class period, which was 4-5 statements for three of the instructors, around 2-3 for one instructor, and between 0 and 1 for the remaining two instructors. In a related study, Turpen and Finkelstein (2010) administered surveys to students to explore the relationship between these instructor differences and student perceptions of their classroom experiences. They found that students who were in a class with more opportunities for discussion felt comfortable asking questions and speaking during class, both with their peers and the instructor. A higher fraction of these students also stated that it was important to understand the reasoning for an answer, not just to know the correct answer.

Teaching Assistants

Graduate and undergraduate students are frequently employed to teach laboratory and/or recitation sections of introductory courses, especially at large institutions. The preparation and instructional support these teaching assistants (TAs) receive varies widely across institutions, and very little research has been done on teaching assistants' beliefs and practices.

- *Teaching Assistants' buy-in to reform instruction* Goertzen, Scherr, and Elby (2009; 2010) used interviews and videotaped observations of teaching assistants to analyze their classroom interactions while using *Tutorials* curriculum at the University of Colorado – Boulder (CU) and the University of Maryland (UM). They found that there was a higher degree of “buy-in” to aspects of the tutorials at CU than at UM. Specifically, TAs at UM did not value the emphasis placed on qualitative reasoning and building intuition. This resulted in TA behaviors that conflicted with the tutorial developers' intentions, such as giving direct answers to students' questions and increasing the use of equations as a reasoning tool. Instead of interpreting these behaviors in a negative light, the authors suggest that TA behaviors can stem from potentially productive attitudes and beliefs about how students learn, and deserve respectful attention (Goertzen, Scherr, & Elby, 2010).
- *The impact of instructional styles on student learning*. Koenig, Endorf, and Braun (2007) studied recitation classes taught by teaching assistants using one unit from the *Tutorials in Introductory Physics* curriculum and four different versions of implementation. The topic studied was energy and momentum, and the four instructional modes were: lecture presentation of the materials, students working individually on the tutorial, students working in a group of 3 or 4 students, and students working in a group with instructor coaching (Socratic dialogue). Student learning in each of these classes was assessed using pre-post test questions designed by the authors, which were written to assess concepts addressed during the tutorial unit. They found that students who experienced the fourth style of instruction scored significantly higher on the post-test than other groups (cooperative groups with instructor coaching). Surprisingly, the students in the third style (cooperative learning without instructor dialogue) scored at the same level as students in

the lecture or individual course sections, suggesting that the TA interactions with students had a substantial influence on learning.

Strengths and Limitations of Research on Attitudes and Beliefs

Strengths.

- There has been at least one large-scale survey research study to identify what fraction of faculty know about PER-based instructional strategies, and how many claim to use them (Dancy & Henderson, 2010; Henderson & Dancy, 2009). This gives a baseline indication of the success attained by dissemination efforts in PER, and where future efforts should be focused. This, as well as interview studies of adoption of PER-informed instructional strategies provides insights into what is needed to accelerate the adoption of effective curricula in teaching undergraduate physics.
- The prevalence of student attitude surveys (like the MPEX and CLASS) has given instructors an easy way to monitor students' attitudes before, during, and after physics instruction.

Limitations

- Qualitative studies involving video and audio-taped interviews or observations are time consuming and difficult to analyze, so there is often a delay in the availability of results from these studies.
- Some attitude surveys, like the VASS and EBAPS, do not have published documentation about the instrument's development and score validation process.
- Research into faculty adoption of reformed curricula is in its infancy stage and a lot more work is needed to understand ways of helping faculty change their attitudes and adopt more effective instructional strategies.

Areas for Future Study

- *Varied implementations of reformed curricula.* This section summarized research studies that described instructors' variations in their implementation of *Peer Instruction*. There is a need for studies that explore how instructors modify and implement other research-

based instructional strategies and materials such as Cooperative Group Problem Solving, Just-in-Time Teaching, or Interactive Lecture Demonstrations. There is also research needed on sociological and infrastructure factors that interfere with adoption of reformed curricula and instructional practices.

- *Teaching Assistants*. There is a need for additional research on the attitudes and beliefs of teaching assistants (TAs), how physics TAs impact the success of instructional reforms, and the implications for professional development of TAs. There is little guidance on how to prepare teaching assistants, both in terms of the components of a teaching assistant orientation program and ongoing professional development opportunities during teaching.

References

Adams, W.K., Perkins, K.K., Podolefsky, N.S., Dubson, M., Finkelstein, N.D., & Wieman, C.E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics – Physics Education Research*, 2 (010101).

Dancy, M., & Henderson, C. (2007). Framework for articulating instructional practices and conceptions. *Physical Review Special Topics – Physics Education Research*, 3(010103).

Dancy, M., & Henderson, C. (2010). Pedagogical practices and instructional change of physics faculty. *American Journal of Physics*, 78(10), 1056-1063.

Elby, A. (1999). Another reason that physics students learn by rote. *American Journal of Physics*, 67(7), S52-S57.

Elby, A. (2001). Helping physics students learn how to learn. *American Journal of Physics, Physics Education Supplement*, 69(7), S54-S64.

Gaffney, J.D.H., Housley Gaffney, A.L., & Beichner, R.J. (2010). Do they see it coming? Using expectancy violation to gauge the success of pedagogical reforms. *Physical Review Special Topics – Physics Education Research*, 6(010102).

Goertzen, R.M., Scherr, R.E., & Elby, A. (2009). Accounting for tutorial teaching assistants' buy-in to reform instruction. *Physical Review Special Topics – Physics Education Research*, 5(020109).

Goertzen, R.M., Scherr, R.E., & Elby, A. (2010). Tutorial teaching assistants in the classroom: Similar teaching behaviors are supported by varied beliefs about teaching and learning. *Physical Review Special Topics – Physics Education Research*, 6(010105).

- Gray, K.E., Adams, W.K., Wieman, C.E., & Perkins, K.K. (2008). Students know what physicists believe, but they don't agree: A study using the CLASS survey *Physical Review Special Topics – Physics Education Research*, 4(020106).
- Halloun, I. (1997). Views about science and physics achievement: The VASS story. In E.F. Redish and J. Rigden (Eds.), *AIP Conference Proceedings Vol. 399: The changing role of physics departments in modern universities* (pp. 605-614). American Institute of Physics.
- Halloun, I., & Hestenes, D. (1998). Interpreting VASS dimensions and profiles for physics students. *Science & Education*, 7(6), 553-577.
- Hammer, D. (1989). Two approaches to learning physics. *The Physics Teacher*, 27, 664-671.
- Hammer, D. (1994). Epistemological beliefs in introductory physics *Cognition & Instruction*, 12(2), 151-183.
- Hammer, D. (1995). Epistemological considerations in teaching introductory physics *Science Education*, 79(4), 393-413.
- Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge, reasoning, and an appropriate role for education research. *American Journal of Physics*, 64, 1316-1325.
- Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics *The Journal of the Learning Sciences*, 12(1), 53-90.
- Henderson, C., & Dancy, M.H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics *Physical Review Special Topics – Physics Education Research*, 3(020102).
- Henderson, C., & Dancy, M.H. (2008). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics*, 76(1), 79-91.
- Henderson, C., & Dancy, M.H. (2009). Impact of physics education research on the teaching of introductory quantitative physics in the United States. *Physical Review Special Topics – Physics Education Research*, 5(020107).
- Henderson, C., Yerushalmi, E., Kuo, V.H., Heller, P., & Heller, K. (2004). Grading student problem solutions: The challenge of sending a consistent message *American Journal of Physics*, 72(2), 164-169.
- Henderson, C., Yerushalmi, E., Kuo, V.H., Heller, K., & Heller, P. (2007). Physics faculty beliefs and values about the teaching and learning of problem solving. II: Procedures for

measurement and analysis. *Physical Review Special Topics – Physics Education Research*, 3(020110).

Hofer, B.K. (2001). Personal epistemology research: Implications for learning and teaching. *Journal of Educational Psychology Review*, 13(4), 353-383.

Hofer, B.K., & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140.

Kortemeyer, G. (2007). Correlations between student discussion behavior, attitudes, and learning. *Physical Review Special Topics – Physics Education Research*, 3(010101).

Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study from introductory physics. *American Journal of Physics*, 73(4), 372-382.

Mason, A., & Singh, C. (2010). Surveying graduate students' attitudes and approaches to problem solving. *Physical Review Special Topics – Physics Education Research*, 6(020124).

May, D.B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics*, 70(12), 1249-1258.

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

Perkins, K.K., Adams, W.K., Pollock, S.J., Finkelstein, N.D., & Wieman, C.E. (2005). Correlating student beliefs with student learning using the Colorado Learning Attitudes about Science Survey. In J. Marx, P. Heron, and S. Franklin (Eds.), *AIP Conference Proceedings Vol. 790: 2004 Physics Education Research Conference* (pp. 61-64). Melville, NY: American Institute of Physics.

Pollock, S.J., & Finkelstein, N.D. (2008). Sustaining educational reforms in introductory physics. *Physical Review Special Topics – Physics Education Research*, 4(010110).

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

Redish, E.F., Saul, J.M., & Steinberg, R.N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.

Turpen, C., & Finkelstein, N.D. (2009). Not all interactive engagement is the same: Variations in physics professors' implementation of *Peer Instruction*. *Physical Review Special Topics – Physics Education Research*, 5(020101).

Turpen, C., & Finkelstein, N.D. (2010). The construction of different classroom norms during Peer Instruction: Students perceive differences. *Physical Review Special Topics – Physics Education Research*, 6(020123).

White, B., Elby, A., Frederiksen, J., & Schwarz, C. (1999). The epistemological beliefs assessment for physical science. Presentation at the American Education Research Association, Montreal (unpublished).

Yerushalmi, E., Cohen, E., Heller, K., Heller, P., & Henderson, C. (2010). Instructors' reasons for choosing problem features in a calculus-based introductory physics course. *Physical Review Special Topics – Physics Education Research*, 6(020108).

Yerushalmi, E., Henderson, C., Heller, K., Heller, P., & Kuo, V. (2007). Physics faculty beliefs and values about the teaching and learning of problem solving. I: Mapping the common core. *Physical Review Special Topics – Physics Education Research*, 3(020109).

Summary & Conclusions

Summary of Research in Physics Education from a Historical Perspective

Even with this non-comprehensive review of PER at the undergraduate level, it is evident that there has been a considerable volume of research done to date. One characterization of PER is that it had its origin in some instructors observing stubborn conceptual difficulties experienced by their students in topics which seemed to the instructors to be simple physics ideas. This interest in digging deeper into students' conceptual learning difficulties led to lengthy catalogs of common misconceptions, which have been documented for nearly every physics topic ranging from mechanics concepts like force, motion, momentum, and energy to topics in electricity and magnetism, thermal physics, light and optics, and modern physics. Research to document students' conceptual difficulties in the 1970s and 1980s led to the development of multiple-choice concept inventories that became widely available in the early 1990s, and to instructional interventions aimed at helping students overcome conceptual difficulties.

The availability of concept inventories served as a catalyst for discussions among physicists about the nature of learning and of conceptual difficulties among their students. Developers of early concept inventories, as well as PER researchers, invited colleagues to administer those inventories to their students following instruction; most refused, considering it to be a waste of time since they thought that surely their students would not hold these incorrect notions about fundamental and easy (to the professors) physics ideas—after all, they were good

instructors who presented ideas clearly. Those few who initially took up the offer were surprised, perhaps even dismayed, by what they found. What later research showed was that quality of lecturing or instructor charisma had little to do with helping students learn concepts about which they held deeply rooted beliefs that contradicted physics laws.

Parallel with research into conceptual difficulties was interest in problem solving, given how central it is to physics. Initial research studies in the late 1970s and early 1980s focused on describing “expert-novice” differences in problem solving, by contrasting the processes used by beginning physics students to the principle-based approaches used by experienced solvers. These studies led to the development of instructional strategies and curricula to promote the use of expert-like approaches, which continue to be active topics of research today. It has also not gone unnoticed with physics instructors that teaching problem solving skills to students in physics is a challenging endeavor. Indeed, nearly all physics instructors have experienced students coming to them and stating that they are A students in nearly all subjects but that they are doing poorly in physics and pleading for some prescriptive guidance for how to do well on exams.

Thus, research on problem solving, combined with research on conceptual change, have given rise to research-based and research-informed concept inventories, curricula, and instructional strategies. Research-based instructional strategies have become collectively referred to as “interactive engagement” methods in contrast to traditional, passive modes of instruction. For example, the use of classroom polling technologies or “clickers” and interactive demonstrations have become relatively widespread in introductory physics courses, and classroom environments structured to support increased interactions are growing in popularity (e.g., workshop, studio, and SCALE-UP classrooms). Several curricular “packages” were identified in the section on *Curriculum & Instruction*, and some of these have been in existence for several decades. In recent years the design of curricula has expanded to include computer-based and web-based instruction, including online homework systems, animations / simulations of phenomena, multimedia presentations, and computer tools for laboratory data collection and analysis.

As some of the more obvious things to try have already been done in PER, the decade from 2000-2010 has seen increased work in more interdisciplinary areas, such as cognition, argumentation in physics from a linguistic or psycholinguistic perspective, and student/faculty

attitudes about teaching and learning, including opportunities and obstacles for instructors' adoption of research-based/informed instructional strategies.

Besides looking at the expansion of PER, another way to characterize the growth and interest in PER is to observe the historical development of the Physics Education Research Conference. Prior to 2001, PER conferences were a “cottage industry” with sporadic conferences held whenever a senior member of the PER community decided to organize one. There was one such conference organized by Bob Beichner at NC State University in 1994 attracting about 43 attendees (24 PER faculty, with the remaining 19 being graduate students and local faculty); the venue was not to discuss research, but issues such as the job market, what graduates of PER programs should know, how they should learn it, what the requirements of a PER graduate program should be, the need for publication venues, future conferences, research funding, and educating the physics community. The next PER conference was organized by Bob Fuller at the University of Nebraska in 1998 attracting 83 participants. Starting in '01, the Physics Education Research Conference (PERC) became a yearly event attached to the end of the American Association of Physics Teacher's summer meeting; a peer-reviewed conference proceeding also became an ancillary part of the conference. Typical attendance at the PERC during the last few years has been over 200 attendees.

What Distinguishes PER from Other DBER Fields?

For the two authors it is hard to characterize PER as one thing since it is diverse and evolving, drawing from disciplines such as cognitive science, education, linguistics, psycholinguistics, assessment and measurement.

It should be noted that there are some elements of physics that distinguish it from other natural sciences, which may have implications for how PER differs from other discipline-based educational research. One is that the content of introductory physics (i.e., classical physics) has changed little in more than a century. This is certainly not the case in the biological sciences, astronomy, and geology, where the content of textbooks from 50 years ago has undergone major transformations compared to the content of today's textbooks; in contrast, other than the addition of color and other cosmetic changes, introductory physics textbooks have changed little. Physics instruction also places a strong emphasis on quantitative problem solving, which is not the case

in other science disciplines (e.g., biology and introductory astronomy). Engineering places a major emphasis in design, while field experiences and a systems approach are very prominent in geology; not so in traditional physics instruction. These differences are likely to have important implications for the type and direction of STEM discipline-based education research.

In summary, we believe PER has a strong research basis in the first four areas addressed in this review: students' difficulties learning physics concepts and solving problems, the development and evaluation of instructional strategies and curricula (particularly "interactive engagement" methods), and the design and analysis of concept inventories. The final two areas of cognition and attitudes and beliefs are less developed but growing. We conclude this synthesis with suggestions and speculations about future directions for PER that could hold promise, or that at least would further inform physics teaching and learning. Although we have done some of this already at the end of the six sections above, here we point to issues and questions that may not fit neatly into one of our six organizational areas.

Future Directions in Physics Education Research

One general direction in need of attention is building theoretical frameworks to guide PER. It likely did not go unnoticed that the Theoretical Framework sections of the six topical areas above were rather uneven, some being highly developed (as was the case for the three competing theoretical frameworks in the Conceptual Change section) and some resembling a pastiche drawn from various disciplines. To develop theoretical frameworks requires researchers and research groups to focus on programs of research lasting many years. However, the funding structure for PER (and for other discipline-based STEM education research) promotes research projects lasting the duration of the grant, with clear beginnings and endings, rather than long-term programs of research where researchers can explore a general theme from multiple perspectives and paradigms that could lead to theoretical breakthroughs. In contrast, research in the natural sciences is much more programmatic, drawing on previous theoretical frameworks in order to refine old theories or build new ones.

Another general area that needs attention is the disaggregation of data in terms of underrepresented minorities or academic majors. Most research studies do not consider multiple, diverse student populations in their design or in reporting results. The small fraction of women

and minorities participating in physics is cause for additional attention to the issues facing these groups, including additional research to explain observed performance differences on concept inventories.

We also see PER moving in the general direction of studying upper-division students (students in advanced undergraduate physics courses and graduate courses). In particular, we expect to see more studies in the near future that identify advanced students' difficulties and adapt research-based instructional strategies and materials for advanced physics courses.

We mention several other areas that could be promising directions for PER:

- *Applications of cognitive load theory to physics learning and teaching.* Cognitive load theory (see Plass, Moreno, & Brunken, 2010, and Sweller, forthcoming, for reviews) is an instructional theory based on what is known about human cognition, in particular drawing on the operation of, and relationship between working memory and long-term memory. Cognitive load theory argues that working memory is limited and overloading it is counterproductive for learning. Many instructional design principles have emerged from cognitive load theory that have proven effective for learning in empirical studies. At present, not a lot of attention has been paid to applying this potentially fruitful theoretical construct in PER. In fact, the next two areas below are also related to cognitive load theory.
- *Learning problem solving by studying worked examples versus solving problems on one's own.* There are research studies in cognitive science cited in the Cognition section indicating that it may be much more efficient to learn problem solving strategies by studying worked examples than by working out problems on one's own. Performing a complicated task such as solving physics problems taxes working memory, and so having students study worked examples (perhaps in combination with solving problems) might be a much more efficient way to promote the acquisition of problem solving skills. We are not aware of research exploring the relative effectiveness of, or interplay between, learning from studying worked examples and learning from solving problems.
- *Learning from text and multimedia.* Textbooks are a multi-million dollar industry yet little research exists on how to best design textbooks to promote students to read and learn from them. There is compelling evidence that the majority of students do not read physics

introductory textbooks, and that multimedia materials designed according to multimedia learning principles are much better than traditional textbooks for learning basic content (Stelzer et al., 2009). The confluence of these findings suggests that a fruitful research direction would be to explore optimal design of learning materials, including form (text only, multimedia, animations, figures, etc.) and function (e.g., reducing memory load). A casual inspection of traditional textbooks reveals flagrant violations of multimedia design principles based on cognitive load theory. Yet, textbooks continue to thrive and add “enhancements” (e.g., color, “boxes”) without evidence of effectiveness. In addition to improving the effectiveness of textbooks and multimedia (e.g., presentations, simulations, animations, online homework, problem solving tutors, computer-based data collection & analysis tools) we expect the evolution of technology will continue to prompt researchers in PER to generate and test new curricular materials that utilize these technologies.

- *The role of motivation, interest, and other affective factors in learning physics.* Although there is ample work on the role of constructs such as intrinsic motivation (Malone, 1981), achievement goals (Elliot & Dweck, 2005), interest (Hidi & Renninger, 2006), task value (Eccles & Wigfield, 2002), affect (Blanchette & Richards, 2010), arousal (Eysenck, 1976), self-efficacy (Bandura, 1997), and attributions (Weiner, 1985) in learning, little of this work has been explored in PER. Applying some of these constructs from social and educational psychology to studies physics learning could hold promise.
- *The role of self-monitoring in physics learning.* The second author of this review, and he is sure all other “veteran” instructors of physics, have had many students come to his office to lament the fact that they studied hard, understood all the material covered in class, did all the homework, and worked out many of the practice exams handed out, only to do miserably on the real course exams. Conversations with those students suggest that they do a poor job monitoring their “test-readiness.” They seem to prepare for exams casually by working out past exam problems with resources at their disposal but never simulating a real test situation in order to monitor readiness. Research from cognitive psychology indicates that students can both overestimate as well as underestimate their level of learning and understanding (Dunning, Heath, & Suls, 2004; Kornell & Bjork, 2009). Findings from this type of research might serve as a starting point for Research on how to help students develop better self-monitoring and test-readiness skills. Further,

research on the design of effective test-readiness tools that students could use to monitor readiness could prove very useful for learning about methods for improving retention of qualified students in the STEM disciplines.

- *Homework and exams.* Although assigning homework and administering exams are basic elements of physics instruction, PER provides very little guidance for instructors on these topics. For homework, it is unclear how much to assign, how to grade it (if at all), what kind of feedback is useful, and whether to assign it on paper or online. Similarly, there is little research to suggest how frequently students should be tested on material, what types of questions are best, and how students should prepare for exams. A promising future direction for PER is studies that apply results from the *testing effect* to science classes (McDaniel, Roediger III, & McDermott, 2007).
- *Learning from laboratory curricula and instruction.* As reviewed in this document, much of the research on physics laboratories has focused on the implementation of computer-based tools for data collection and analysis. Only a few published curricula exist for laboratories (e.g., Investigative Science Learning Environments, RealTime Physics Active Learning Laboratories) and it remains largely undocumented in the literature what other lab resources are used in physics courses. There is a need for research on what students learn by participating in laboratories.

In conclusion, Physics Education Research as a discipline has a rich history, makes a wide array of interdisciplinary connections, and yet has many promising avenues for future research.

References

- Bandura, A. (1997). *Self-efficacy: The exercise of self-control*. New York: Freeman.
- Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making, and reasoning. *Cognition and Emotion*, 24, 561-595.
- Dunning, D., Heath, C. & Suls, J.M. (2004). Flawed self-assessment: Implications for Health, Education, and the Workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.
- Eccles, J.S., & Wigfield, A. (2002). *Development of achievement motivation*. San Diego, CA: Academic Press.

- Elliot, A.J. & Dweck, C.S. (2005). Competence and motivation: Competence as the core of achievement motivation. In A.J. Elliot, C.S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 3-14). New York, NY: The Guilford Press.
- Eysenck, M.W. (1976). Arousal, learning, and memory *Psychological Bulletin*, 83, 389-404.
- Hidi, S., & Renninger, K.A. (2006). The four-phase model of interest development *Educational Psychologist*, 41, 111-127.
- Kornell, N. & Bjork, R.A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning *Journal of Experimental Psychology: General*, 138(4), 449-468.
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction *Cognitive Science*, 4, 333-369.
- McDaniel, M.A., Roediger III, H.L., & McDermott, K.B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14(2), 200-206.
- Plass, J.L., Moreno, R., & Brunken, R. (Eds.) (2010). *Cognitive Load Theory*. New York, NY: Cambridge University Press.
- Sweller, J. (forthcoming). Cognitive load theory. In J.P. Mestre and B.H. Ross (Eds.), *The Psychology of Learning and Motivation: Cognition in Education*. San Diego, CA: Elsevier / Academic Press.
- Stelzer, T., Gladding, G., Mestre, J.P., & Brookes, D.T. (2009). Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content. *American Journal of Physics*, 77(2), 184-190.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.