Identifying and Addressing Difficulties:
Reflections on the empirical and theoretical basis of an influential approach to improving physics education

Paula R. L. Heron

University of Washington, Seattle WA 98195-1560

Abstract:

The past few decades have seen dramatic growth in Physics Education Research (PER). Along with this growth, there have been significant shifts with respect to which questions to pursue, how to pursue them, and even how to talk about findings. Throughout this evolution, the Physics Education Group at the University of Washington has conducted an influential program of research and curriculum development and maintained a focus on conceptual understanding and reasoning skills. One tradition closely associated with the group is improving student learning by *identifying and addressing student difficulties*. This paper attempts to set forth the empirical and theoretical basis of this tradition, to place it in the broader context of PER, and to outline an agenda for future research that represents a natural outgrowth of the empirical achievements of the past in the light of emerging theoretical developments.

1. Introduction

The Physics Education Group at the University of Washington (UWPEG) is one of the most enduring institutions in physics education research (PER). The group, founded by Lillian McDermott in the 1970s, has had a significant influence on intellectual and cultural traditions of the field, including the topics and methods of study, the principles of graduate education, and the styles and venues for presenting findings. Therefore, a closer examination of the group's beliefs, values, and practices, which have rarely been explicitly discussed in our publications, may be of some value in understanding how PER arrived at its current state, and how it might continue to evolve.

One tradition closely associated with the UWPEG is improving student learning by *identifying and addressing student difficulties*. This paper attempts to set forth the empirical and theoretical basis of this tradition. However, while I have been a part of the UWPEG for more than twenty years (first as a postdoc, now as a professor), I do not claim to speak for the entire group. Therefore, throughout this essay I have mingled the use of "I" and "we;" the former usually signals a point on which I suspect my colleagues might not agree precisely. Some parts of this paper are based on my recollections. While I refreshed my memory by consulting articles, conference agendas, etc.*,* the historical notes should not be taken as resulting from a scholarly study – rather they are my impressions, influenced by my own preoccupations at the time and altered, inevitably, by the passage of time. Nevertheless, I believe that to the extent they help explain the present, my reflections may prove helpful.

My main goal is to explain how *identifying and addressing difficulties* fits within the broader context of PER. Along the way I will try to address common questions raised about "difficulties," and to examine how my views have evolved. In particular I will explicitly discuss theory: education theories in general, specific theories we in the UWPEG have and have not embraced, and why the status of theory in our group seems to be an enduring point of interest for PER.

I think of a PER program as reflecting a researcher's (1) values – what is important to them in physics education; (2) goals – what they want to achieve; (3) beliefs about learning – how they think learning happens; (4) customary practices – what they believe are effective procedures; and (5) context – what affordances and constraints their research setting offers. For example, it may be that a researcher thinks that developing a sophisticated epistemology for physics is important and is dissatisfied with the current state of affairs in her classroom, or in the discipline as a whole. She therefore sets out to design a course that emphasizes epistemology. To do this she must have a set of guiding beliefs or principles – for instance she might consider it important first to

understand what students' existing ideas are, and how those ideas came to be. She might consider it vital to understand how experts came to their views. She might believe that students will have a supply of ideas that represent the raw material from which more sophisticated views can be refined. To explore this raw material she might conduct interviews to gain deep insight or develop multiple-choice tests to survey large groups. While this example represents a very applied research program, I think a parallel set of decisions underlies more pure or fundamental research as well.

To some extent the organization of the main sections of this paper follows this structure. First, in Section 2, I provide some broad context by reviewing some developments in PER over the past few decades.[i] Section 3 provides some more specific context about the UWPEG and *Tutorials in Introductory Physics.*[1] Section 4 focuses on the instructional values and goals that help shape our program, and provides an overview of how we employ *intentional teaching* to try to improve student learning. Section 5 describes the perspective of *pragmatic constructivism* and the specific assumptions about learning that help guide our efforts within the framework set forth in Section 4. Section 6 discusses some practices we follow to assess the effectiveness of individual tutorials and determine if any increases in performance are *instructionally significant*. Finally, there is a short discussion of some current research directions that are motivated by puzzling observations for which we have not (yet) found satisfactory explanations.

2. PER: Reflections on the past two decades

Research on the learning and teaching of physics has a long history.[2-4] Rather than trying to provide a summary here, or a comprehensive overview of the current state of the field, I will try to sketch some of the most relevant elements for understanding the context in which the *Tutorials* were developed, and in which the views about learning and teaching I express here were formed.

In the mid-1990s, I joined the field of PER after obtaining a PhD in condensed matter physics. At that time, conceptual understanding was a dominant theme. The bulk of the research concerned introductory university courses; research in more advanced physics courses was just emerging. Conference presentations and publications were dominated by the promotion of improved learning through various interactive engagement strategies[5] and assessment through multiple-choice concept inventories. Local, regional

---

[i] The focus is almost entirely on PER is the USA, for the sole reason that debates and developments in the US-based community are most relevant to the specific issues I address in this article.

and national meetings featured talks and workshops on active learning approaches including *Peer Instruction*,[6] *Workshop Physics*,[7] *Interactive Lecture Demonstrations*,[8] *Real-Time Physics* labs,[9] *ALPS* worksheets,[10] and *Cooperative Group Problem-solving*.[11] In a development that seemed revolutionary at the time, the authors of these materials offered evidence of their impact on student understanding, not just reports of student satisfaction.

Quantitative problem-solving was perhaps the other pillar of PER, with several groups devoting significant effort to characterizing and bridging the expert-novice gap.[12,11] Social and affective aspects of physics teaching and learning were not the subject of a significant amount of study. What little there was focused on the interaction with conceptual learning, such as documenting gender differences on *Force Concept Inventory* (FCI) scores.[13,14] This is not to say that the PER community was not engaged in thinking about broadening participation, just that there were relatively few conference presentations and even fewer published papers. Affective issues and students' views about how best to learn physics were beginning to receive attention. The publication of the *Maryland Physics Expectations Survey* (MPEX) in 1998 marked a turning point for these studies, which have since become a mainstay of PER.[15]

The 1999 Resource Letter on PER by Redish and McDermott in the *American Journal of Physics* provides a good snapshot of the era.[16] Of its 224 entries, 115 are devoted to conceptual understanding, 12 to problem solving and 11 to attitudes and beliefs. The *Proceedings of the 1998 Physics Education Research Conference* in Lincoln, NB are another window into the concerns of the field in those years.[17] One interesting section is the transcript of a panel on "Frontiers in Research in Physics Education." Panel participants, (Jose Mestre, Ron Thornton, Bruce Sherwood, and Alan Van Heuvelen) called for: the development of more assessment instruments; rethinking *what* we teach, not just *how* we teach; additional work on representations; schemes for "diagnosing" individual students to customize learning experiences; and developing workplace skills. One participant noted that "many people are now starting to develop" simulations and animations and that this should be a fruitful area for research and development for the development of conceptual understanding. Some of the calls in that session proved prescient. However, there has been growth in what we study in PER and how we study it that far outpaces any predictions made in the 1990s.

The subsequent development of PER is well captured by the annual PERC Proceedings. In 2003, after self-publishing peer-reviewed proceedings for a few years, PERC organizers made arrangements for publication through the American Institute of Physics (AIP). (They are now self-published again, on Compadre.org.) From 2003 through 2012, the Proceedings grew significantly in size and scope. The 2003 volume

contains 40 peer-reviewed articles touching on issues of gender and self-efficacy as well as the learning of content.[18] Authors drew on theoretical frameworks including coordination classes, mental models and resources and new (to PER) methods, such as gesture analysis. By 2012, the number of peer-reviewed papers had grown to 110.[19] Gender, identity and self-efficacy were the subject of several articles, using participationist and intersectional perspectives, among others. Content learning was viewed from multiple perspectives such as embodied cognition, discourse analysis, visual perception, conceptual blending, metacognition and cognitive linguistics. Discourse analysis and social network analysis were used to analyze the student experience.

By now the field has broadened to the extent that the categories in the Redish & McDermott *Resource Letter* are no longer adequate for capturing the scope of scholarly efforts aimed at understanding and affecting how people learn, teach, participate, persist, and succeed in physics. Table I contains a non-exhaustive list of popular research themes associated with three major domains. The *intellectual* domain includes research on the how learners develop and apply ideas about the natural world and about the discipline of physics. The *affective and individual* domain includes research on the emotional experiences of learners and their views about themselves in relation to the discipline. The *social and cultural* domain includes research on the classroom environment, the culture of physics, and the influence of the broader culture in which learners operate. While these three domains can serve as a map to the PER landscape it must be noted that there are many other ways in which one could represent the field; this scheme simply highlights different aspects of learners' experiences that we, as teachers and researchers, might take into account and/or be able to influence.

Table I: Some current topics in PER

The intellectual domain
conceptual understanding
reasoning, mental models, metaphors, framing
spatial-visual reasoning
quantitative problem-solving
measurement and experimentation
"thinking like a physicist"
mathematical thinking in physics contexts
the epistemology of physics (and the nature of science more broadly)
ideas about the nature of teaching and learning
The affective and personal domain
self-efficacy and identity
attitudes, expectations and personal epistemology
The social and cultural domain
participation of underrepresented groups (women, people of color, individuals who identify as LGBTQ, disabled persons)
social networks, group dynamics
classroom and institutional environments

For each of these themes there are associated efforts to understand structures, relationships and mechanisms, to develop different tools, strategies, and environments, and to prepare instructors for their roles. I believe that most physics education researchers value insights into these topics for their own sake *and* for the potential they offer for improving student learning, deepening engagement and broadening participation. However, some researchers focus primarily on understanding processes, and are only incidentally concerned with whether their results lead to changes in how physics is taught. For others, changing how physics is taught is the primary motivation; research serves to inform and validate their efforts.

I should acknowledge that there are several ways in which the list in Table I fails to capture the true nature of PER. One is that it does not convey the degree to which a lot of PER work (especially emerging work) takes place where different areas overlap. For instance, investigating the impact of participation in social networks on achievement bridges the social and intellectual domains.[20] Another shortcoming is that by focusing on *what* we study, there is no acknowledgement of the increasing diversity in research methods. In particular, the growth of rigorous qualitative research methods has significantly changed the face of PER.[21]

The third way this list falls short is related to the two roles that the term 'physics' serves in the phrase 'physics education research.' One role is to convey that the subject of research pertains to the learning and teaching of physics; the other is to convey that the manner in which research is conducted is derived from, or at least reflects, the manner in which physics research is conducted. This view is somewhat controversial. It can be argued that research in PER is (or should be) based on established procedures in the social sciences, such as randomized control trials of instructional interventions. However, these procedures are themselves based to a large extent on idealizations of classical physics research, characterized by hypothesis testing through controlled experiments. PER in the US was historically dominated by researchers whose only formal training was in physics[ii] and who entered into research on learning and teaching with expectations about the complexity of experiments, measurement uncertainty, the importance of replication, the nature of theories, etc.[22] They also carried with them cultural expectations about the nature of graduate training and the beneficial role of competition among many researchers working toward the same goals. Many physics education researchers continue to identify strongly with physics and feel most at home in physics departments and at physics conferences. However, as PER has grown, an increasing number of researchers are earning PhDs in schools of education and are most at home at meetings of researchers in science education, cognitive studies, or the learning sciences. Still others are most at home in multi-disciplinary institutes or centers. As the field evolves, the norms, values and practices of different scholarly traditions are increasingly apparent, and often in tension. The extent to which physics education research should continue to be influenced by the formal physics training of researchers is an emerging debate for the field. It is also no longer possible for any single researcher to be fully versed in all areas of the field, as it once was. Thus, it may turn out that PER splits into sub-fields, with significantly less interaction, much the way that condensed matter physics and particle physics have their own dedicated journals, conferences, etc.

3. The UWPEG and *Tutorials in Introductory Physics*

Early papers by Lillian McDermott and her students focused on concept development among preservice and inservice K12 teachers, and university students aspiring to STEM careers who lacked the background to succeed in "gateway" courses like introductory calculus-based physics.[23-25] While PER has broadened significantly, the UWPEG has maintained its focus on physics-specific concepts and reasoning skills and

---

[ii] Examples in the US include Robert Fuller, Fred Goldberg, Ken Heller, Priscilla Laws, Lillian McDermott, José Mestre, E.F. (Joe) Redish, Fred Reif, David Sokoloff, Ron Thornton, Alan Van Heuvelen, and Dean Zollman

has continued to develop instructional materials that require minimal equipment and little or no technology. However, over the years, the scope of research has broadened to include introductory physics, upper-division physics, and to encompass learning in lectures, labs, tutorials and intensive courses for K12 teachers. Publications in *The Physics Teacher* and the *American Journal of Physics* as well as talks and workshops at meetings of the American Association of Physics Teachers, have allowed group members to reach practitioners – physics educators at all levels – who could use the results of research in their own classrooms.

When I arrived at UW, the group was producing a preliminary version of *Tutorials in Introductory Physics.*[26] Tutorials had been in use in all sections of the UW introductory calculus-based physics course since 1991. By 1995, several thousand students had participated and a few dozen graduate students had gone through the associated teaching assistant (TA) preparation course. The tutorials were intended to supplement instruction by lecture, textbook and laboratory, and in many implementations, replaced a standard recitation or problem-solving session. Like the "active learning" innovations mentioned earlier, they were meant to engage students in thinking more deeply about the topics covered in nearly all introductory college and university courses.

Since publication of the preliminary edition in 1998, tens of thousands of students have participated in courses featuring the UW tutorials at dozens of colleges and universities in the US and abroad, both in the original English and in translation. The tutorials proved to be a fertile ground for research at UW and elsewhere. In addition to numerous citations, in one recent five-year period (2011 – 2015), more than 100 articles, presentations and dissertations focused on the implementation and assessment of tutorials at other institutions.[27,28] This list continues to grow.

Given the importance of the tutorials to the PER community as a vehicle for teaching, research, and the preparation of TAs and future teachers, it seems appropriate to examine some of these issues more explicitly than has been done in the past. I should note that to a large degree the discussion below applies equally to *Physics by Inquiry,*[29] but for the sake of simplicity, I will limit my focus to the tutorials.

4. Instructional goals and values: "Intentional teaching"

The first tutorial I taught had to do with velocity as a ratio of a distance traveled[iii] to the time interval over which the travel took place. It directly addressed the concept of

---

[iii] The motion was strictly in one direction and one dimension, therefore the distance traveled and displacement are identical.

instantaneous velocity as the ratio over short time intervals. Each student was given a short piece of ticker tape: every piece had the same number of dots, but the lengths of the pieces varied. Each student calculated the (average) velocity represented by their own piece of tape. They assumed, in most cases implicitly, that using a single number to represent the velocity was unproblematic and indeed the dots appeared evenly spaced. The entire group of 22 – 24 students was then invited to arrange themselves from highest to lowest velocity. With some prodding, they recognized that the pieces of tape fit together and in fact had been obtained from a single object travelling with *changing* speed. This exercise was thrilling to me (less so, I must admit, to my students). I had never before contemplated instantaneous velocity in these terms. To this day, this tutorial represents to me an ideal of sorts. Difficult concepts were related to a real-world motion. Ratios were interpreted verbally. Limits were tackled directly. The idea of instantaneous velocity was introduced before the term.

The *Velocity* tutorial, like all other tutorials, was intended to help students develop a functional understanding of a major topic in a typical introductory course. PER studies have revealed the deficiency of many standard textbook presentations. For example, flaws in treatments of the work-energy theorem have been thoroughly documented.[30-32] However, we have not typically advocated for changing what is taught in introductory physics, other than recommending a general reduction in content. Instead we have taken the typical curriculum as a starting point and asked what we could do to improve how it is taught, and in turn, what students learn. Thus our investigations typically (though not exclusively) aim to reveal what students can and cannot do after other forms of instruction in the course, namely lectures, labs, textbook readings and homework assignments, have been completed. Specifically, we ask: what do we want students to be able to do? To what extent are these goals are being met? In what ways are we falling short? What are the barriers to meeting those goals? What do we have to work with – in other words, what are students doing well?

In courses at UW, most students learn to apply rules and procedures with sufficient mechanical competence to obtain passing grades; what they often cannot do is reason qualitatively with concepts and principles. Their efforts often reveal a failure to distinguish closely related quantities (velocity and acceleration, potential and potential difference, kinetic energy and momentum, heat and temperature, etc.). Students display uneven skill with representations (vectors, graphs, field line diagrams, ray diagrams, etc.). In reasoning with multivariable relationships ($PV = NkT$, for example), they frequently ignore (perhaps un-knowingly) one variable. There is a widespread tendency to think locally about changes to a physical set-up. Students often fail to grasp the implications of critical simplifications such as frictionless surfaces, massless strings,

ideal capacitors, perfect thermal insulation, etc. Above all, many students lack models with which they can reason when formulas are inadequate.

Each tutorial has been composed to help "bridge the gap" between what students are taught through conventional instruction, and what they actually learn.[33] They take students' ideas into account in the broadest sense. I think of the approach as ***intentional teaching:*** each step in an instructional sequence is carefully chosen with specific goals in mind and with a view to the likely thought processes of the learner. The design of exercises, questions and experiments is highly deliberate. A tutorial does not provide a rigid pathway from which no deviation is possible, nor are the step sizes so small that reasoning sequences can be followed unerringly. The instructors (typically graduate or undergraduate teaching assistants) must exercise judgment and creativity in responding as groups tackle thorny issues: When to ask an additional question to sharpen the issue? When to ensure that all viewpoints have been properly heard? When to step back and allow conversation to proceed organically? The result is that no group, or individual, navigates a tutorial in exactly the same way. While students' own interests do not drive the curriculum (the topics are predetermined), it is taken for granted that students are sincere actors, who want to learn and who are capable of reasoning that has often been thought to be beyond their capabilities.

Many tutorials provide direct assistance to students in interpreting common representations such as vectors, graphs, free-body diagrams, field lines, ray diagrams, phasors, etc. Others provide guidance with broad themes such as choosing systems, considering limits (such as treating a thin string as the limiting case of a rope as the mass goes to zero), applying the principle of superposition, and modeling continuous distributions as sets of discrete elements (as in the case of a line of charge modeled as a distribution of many closely-spaced point charges).

The tutorials take into account research findings about the topics in question. A series of tutorials that helps students construct a model for physical optics provides an example.[34,35] Research revealed that students did not have a clear understanding of how path length differences give rise to phase differences or how a finite aperture can be modeled as a series of point sources. As a result, they had little ability to reason about interference and diffraction effects when the few formulas they had learned were inadequate. The series starts with an exploration of circular wavefronts produced by two sources in a tank of water. Students are led to recognize that the disturbance at different points can be traced to the fact that the wave arriving from one source will have traveled a different distance than the wave arriving from the other. They generate rules for determining the numbers of nodal lines and lines of maximum constructive interference. Subsequent tutorials make an analogy with light. The double-slit

experiment gives way to multiple very-narrow slits, which in turn provides the basis for modeling a finite aperture as the limiting case of infinitely many slits of infinitesimal spacing. Students emerge with a model that they can use flexibly to predict and account for interference and diffraction effects from arbitrary arrangements.

In many cases, tutorials deliberately attempt to build on what students know, or expect to be true. For example, a tutorial on angular momentum leverages students' strong expectations that certain quantities will be conserved.[36] Specifically, we found that many students made incorrect predictions about the results of collisions between point particles and rigid objects, predictions that *would* be correct if linear momentum could be transformed into angular momentum – in other words if the two types of momentum were jointly, rather than separately, conserved. The tutorial relies on students being relatively unwilling to accept that quantities of motion are not conserved in the collisions, preferring to accept that a point particle moving in a straight cam have angular momentum.

As in the illustrations above, in most tutorials, students' attention is deliberately drawn to situations in which research has revealed a tendency to falter. For example, in a tutorial on 2-D motion, not only do students construct the concept of acceleration, they are led to consider motion on a variety of closed paths, because of a well-documented tendency to inappropriately generalize the result for circular paths and assume that acceleration is directed toward the "center" of non-circular paths.[37] In the sequence on physical optics, students' attention is drawn to implications they might not have otherwise appreciated. For example, the fact that if the width of a single slit is less than the wavelength of the light passing through it, no areas of complete destructive interference can be seen. Thus, most tutorials help students construct, interpret and apply fundamental concepts and principles, while at the same time addressing "specific difficulties," as discussed below. However, a coherent conceptual framework that enables students to interpret physical scenarios, choose and apply appropriate analysis tools, and evaluate their conclusions, cannot be developed by a piecemeal approach that treats difficulties in isolation.

5. "Pragmatic constructivism"

Despite the hundreds of talks, workshops and articles we have produced, there continues to be some lack of clarity in PER about the guiding principles behind the development and assessment of the tutorials. To be sure, we have been reticent about discussing our work in "theoretical" terms, and so our views must be inferred by the reader. As a result, I believe that many in the PER community continue to associate the tutorials with some of the views about "misconceptions" and active learning that

prevailed in the mid-1990s, when the tutorials first emerged. In fact, there is a perception in some quarters that the UWPEG is stubbornly clinging to an outmoded theory (the "theory theory") but refusing to acknowledge it. Our reticence in discussing theory stems in part from a culture we share with our departmental colleagues: the culture of physics. For traditionally trained physicists, many education theories seem more like prescriptions of best practices or summaries of patterns in observations, rather than precise, broad, and powerful explanations invoking unseen entities, interactions and processes. Many of our colleagues express skepticism about the validity of theories in essentially *all* domains outside of physics, where operationally defined concepts and quantitative models seem elusive, even impossible. Therefore, even the general notion of education theory can meet resistance in our context. Yet, for us, and for a great many others in the PER community, physicists in traditional areas are an important audience for the output of our scholarly efforts and it is critical that they find our findings trustworthy. I believe that intellectual honesty nonetheless compels acknowledgment of any theoretical commitments. The absence of such discussions therefore cannot be accounted for solely by cultural discomfort or political calculations. One reason for our relative silence is ambivalence about the status of the broad conceptual framework (discussed in greater detail below) that has provided us some loose guidance. Another reason is that many of the more detailed theories of learning proposed in the past few decades are not broad enough to encompass the spectrum of intellectual challenges in teaching physics and we have been wary of embracing frameworks that might too narrowly constrain our vision. Nevertheless, there are principles, some of which are rooted in unproven assumptions about how the mind works (i.e., they are "theoretical"), that guide our work.

When the tutorials were first developed, the view that students hold persistent "misconceptions" was widespread in the PER community (though perhaps not yet in the broader physics community). In a range of contexts (mechanics, geometric optics, heat and temperature, etc.), students had been observed to respond to physics questions in ways that seemed sensible in the light of everyday experience, but were nevertheless incorrect. Many students clearly articulated and vigorously defended their views. Responses before and after formal instruction were often indistinguishable.[38] A classic example is the notion that an object in motion must be acted on by a (net) force in the direction of that motion such that a coin tossed upwards is assumed to be subject to an upward force while it travels upward, no force at the top of its trajectory, and a downward force while it falls.[39] It was commonplace to hear students referred to as "Aristotelian" as if they were arrested at a similar stage as earlier scientists and philosophers.[40] Much of the literature of the early-1980s to mid-1990s assumed, explicitly or implicitly, that such a belief would be expressed more-or-less reliably across situations involving motion. The prevailing view was that these misconceptions

(also referred to as *preconceptions, naïve beliefs* or *alternative frameworks[iv]*) had been constructed by students on the basis of their experience of the natural world.[39,41] Teachers were warned that students did not arrive in our classrooms as blank slates. I think of this as the era of the "three C's" – constructivism, conceptual change, and cognitive conflict. The term "constructivism" had been adopted and used to differentiate this model of learning from the implicit *transmissionist* view of conventional instruction.[42,43] This perspective also offered an explanation for the apparent failure of conventional instruction: once formed, and found to be satisfactory, it was believed that students' common-sense that conceptions were resistant to change. It was further assumed that students must be given *proof* of the inadequacy of their previously developed ideas.[44] For example, it was understood that they needed to be shown that no force is needed to maintain uniform motion.[45] Various "cognitive conflict" strategies were in vogue as instructors sought to induce conceptual change in their students.[46]

Not everyone subscribed fully to the views outlined above. The diversity of challenges to teaching and learning did not seem to be accountable for by a single theory, at least not an available theory of conceptual change. In particular, some researchers (McDermott, later Shaffer, myself and other collaborators), challenged the notion of stability and coherence:[47]

> *"The students' responses, both in word and action, indicated that they lacked a consistent conceptual system. Their use of the word "force" and other technical terms was ambiguous and unstable. The comments students made during the interview and their readiness to revise unsuccessful strategies made it clear the underlying conceptual problems were complex and could not be adequately summarized as a simple belief in the necessity of a force in the direction of motion." – McDermott, 1984*

Other researchers who recognized the shortcomings of the "theory theory," included diSessa,[48,49] and later Redish, Hammer, Scherr and their co-workers.[50,51] They pushed for greater theoretical clarity and attempted themselves to work out mechanistic accounts of physics learning. An example is the "knowledge in pieces" viewpoint espoused by diSessa. This view did not gain traction immediately, perhaps (and this is speculation) because it was not promoted heavily within the physics community: in fact diSessa's seminal (and very long) paper on the subject was published in *Cognition and*

[iv] It should be noted that within the PER community these terms, as well as *mental models* and *schema,* were used somewhat loosely.

Instruction, not a physics journal. Nevertheless, articulating an explicit theoretical framework became popular in PER.

In spite of the rising expectation of explicit attention to theoretical commitments, we in the UWPEG did not often speak or write about the conceptual framework in which we operated. When we failed to claim our own theory, one was attributed to us. The term *difficulty* was taken to be a synonym for misconception, a term it was assumed we avoided because of its negative connotations. Instead, as mentioned earlier, we were cautious about using terms our colleagues in traditional physics fields would find overblown, even ridiculous; we were cognizant that there was much more going on than "misconceptions;" and we were focused strongly on practical applications. While we have always been sensitive to regularities that occur across topic areas, we have been prepared to act in spite of a lack insight into underlying mechanisms. The following quote describes our perspective well:[52]

> *"We ought to engage in research endeavours which will uncover student frameworks, investigate the ways they interact with instructional experiences, and utilise this knowledge in the development of teaching programmes. It should be emphasized, however, that we are not claiming that this sort of work is atheoretical. Clearly it is not, and cannot be, as some form of theoretical commitment or speculation will always be present in any form of empirical inquiry. ... we must make certain assumptions about the nature of knowledge and the nature of the learner which establish the frame for interpreting and manipulating the teaching/learning environment. ... The important point is whether the frame itself is the subject of inquiry. In the problem-oriented research programme the primary focus is the development of instructional practices which improve science teaching, as opposed to making a claim about the model or theory being used." – Driver & Erickson, (1983)*

During the late 1990s and 2000s I was often asked to explain the UW view.[53,54] Over the years my own views have continued to evolve and my own hesitation in discussing learning theories has abated. The PER community seems to have developed a more sophisticated notion of theory in general, perhaps in part because of the increasing diversity of research interests and the increasing recognition of the complexity of the individual and social systems we are examining, as well as the growth in the sophistication of research methods.

5.1 General Conceptual Framework

The perspective on learning that prevails in the UWPEG can be called **pragmatic constructivism.** It is *pragmatic* in that we address problems that arise in authentic classroom teaching and learning situations. It is *constructivist* in that we assume that learners construct new knowledge on the basis of their existing knowledge (whether canonical or not), which in turn has been constructed on the basis of prior formal instruction and everyday experience. Prior knowledge is viewed both as the foundation upon which new knowledge is built, as well as the building material. We make no *a priori* assumptions about the structure or stability of learners' initial ideas. They may be more stable, elaborated and interconnected in some cases than in others. Thus, we do not view all conceptual difficulties as stemming from firmly held erroneous beliefs; some may stem from the deployment of finer-grained cognitive resources in situations in which they do not apply. Moreover, we recognize that learning takes place in a social environment in which even the act of taking a test is shaped by social norms and expectations, and that it is influenced by affect and identity. Finally, we believe that instruction must acknowledge learners' reasoning strategies, which may have proven adequate in everyday life, but which may not meet the standards of formal, model-based reasoning.

To the extent that the UW PEG uses an over-arching instructional principle consistent with this perspective, it is to *teach by questioning* in an effort to guide and support students in the construction of understanding – in effect to help them identify the pieces and see how to put them together. Therefore, rather than merely tell students what we wish them to learn, we attempt to draw from them the meaning of concepts and relationships. This does not mean that we expect them to discover the laws and theories of physics in one hour each week. In most cases, tutorials assume that a given idea has been presented previously (i.e.*,* that momentum is defined as the product of mass and velocity, and that is it conserved) but that what it *means* to be conserved can be worked out by answering a series of questions. This approach ensures that students go through the relevant reasoning themselves, rather than attempting to follow while someone else (instructor, textbook author, etc.,) outlines that reasoning. Listening to or reading explanations leaves many learners unaware of the steps that were skipped, the choices that were made and the counter-arguments that may arise. Questioning also serves to elicit students' own intuitions and prior knowledge, and to promote reflection on what they do and do not understand.

5.2 Characteristics of "difficulties"

The concepts and principles of physics are not obvious or intuitive and the process of constructing understanding is not smooth, linear and trouble-free. However, there are some predictable ways in which many learners will proceed – some of which lead to errors with important implications. In such cases, we in the UWPEG use the term *difficulty,* which I attempted to define in 2003 as follows:[54]

> *"[A difficulty is] the use of a specific idea or pattern of reasoning instead of those we consider correct and appropriate. The notion that "closer means stronger" might be a value-neutral resource. The assumption that the brightness of lightbulbs in a circuit depends on their proximity to the battery is a difficulty." – Heron (2003)*

It is important to emphasize that in this definition, the difficulty is not the specific idea or reasoning pattern, it is the *use,* or *misuse,* thereof. It is also important to emphasize that "difficulty" is an umbrella term that encompasses a broad range of phenomena. It is not a theoretical entity in the sense of diSessa's phenomenological primitives (*p-prim)*.[48] That is, there is no single hypothetical mental structure or process represented by the term "difficulty." Both misconceptions and *p-prims* are often thought of as "things" that exist in the mind, waiting to be called up when needed. For example, students may be said to "hold" misconceptions or to "activate" a *p-prim*. Similarly, students are often said to "have," or to "encounter," a difficulty – suggesting they are object-like. However, we also sometimes say that students "experience" a difficulty – suggesting an ontological status more akin to a state or a process. If responses to physics tasks are the results of (implicit or deliberate) reasoning and decision-making processes, then difficulties may be thought of as flawed processes: *choosing* an inappropriate or incorrect rule, *ignoring* an important variable, *failing* to recognize a conflict, *blending* similar concepts; *making* an unwarranted assumption, *drawing* an invalid inference.

Difficulties can be thought of as arising from learners' attempts to interpret the content of their formal physics instruction in terms of what they already know. Difficulties are not so much the *content* of that prior knowledge – ideas or beliefs that are called into play when a learner must generate a response to a question – but they arise at the intersection of prior (or emerging) understanding and new rules, concepts, or models. For example, our group documented a tendency to interpret the wave properties of matter in terms of particles moving along sinusoidal trajectories.[34] While this is not the model we wish students to adopt, it must be recognized that it is a *rational* attempt to link prior knowledge with newly introduced ideas. Moreover, we don't assume that

students constructed this view by synthesizing their experience of the real world. Presumably the sinusoidal trajectory idea does not predate formal instruction but instead came into being at the moment it was needed.

(I should acknowledge an underlying assumption in the discussions throughout this article is that there are, in fact, object-like entities in the mind that can be built, stored, retrieved, modified, etc. For a long time, it did not occur to me that there could be any other theories of mind – perhaps a result of my background in traditional physics.[55])

The view sketched here is consistent with the *resources theory,* which explains thinking as drawing on ideas that may be "smaller" in some sense than misconceptions and less tightly linked to physics terms.[51] Resources may be rough precursors to more canonically correct physics ideas, needing refinement perhaps, but valuable as starting points. For instance, the tendency to associate an unbalanced force with an increase in speed may be a step towards understanding Newton's second law.[56] It is important to emphasize that resources and difficulties do not correspond to viewing student thinking from a "glass half full" rather than a "glass half empty" perspective. Difficulties emphasize the points where instructional intervention is called for; resources emphasize the tools that instructors can draw on for those interventions.

Interpreting student thinking in terms of difficulties has some advantages. As I noted in 2003, the term is "broad and inclusive as it comprises both topic-specific and topic-independent ideas and reasoning patterns."[53] The context that accompanies a description is useful. For example, claiming that students tend to assume that current is used up in a DC circuit tells the reader something about where and how the difficulty manifests. Such a description also provides some practical guidance in that it suggests an instructional objective, although not necessarily an instructional strategy. As stated above, multiple models of knowledge and learning can be accommodated under the umbrella of "difficulties." In particular, the existence of misconceptions, if taken to be robust, well-articulated, firmly held beliefs, can be accommodated. So can the existence of p-prims. Therefore, it is the case that "*difficulties"* is a theoretically ambiguous term. Not because of a reluctance to name the theory, but because different theories are required to explain different types of difficulty, as the examples below illustrate.

It should also be noted that there are a number of things we don't assume about learning. For one, we don't assume that all students who make the same error make it for the same reason, or that what a student says or writes in response to the prompt "Explain your reasoning" is an accurate representation of how that student actually arrived at his or her answer (I will have more to say about this later). We also don't assume that students experience difficulties as difficulties – that is that they experience

confusion, frustration or uncertainty. We don't assume that the ability to answer a given question on a topic implies the ability to answer all related questions.

5.3   Examples of difficulties

There are many types of difficulty, illustrated by the examples below. As can be seen, difficulties are inherently contextual, at least up to a point. However, without links to broader patterns, they do not provide an especially satisfying or generative framework. It should also be noted that nearly every response can be interpreted in several different ways – in fact it is a challenge to arrive at a description that both honors the context but is not so specific that it seems unlinked to any broader framework that can allow patterns to be seen, and perhaps explained, in more general terms. In many cases, we tend to summarize our observations, i.e., we give "operational" descriptions. In others we speculate a bit more about the thinking processes that would account for the observations, i.e., we give "theoretical" descriptions, as defined by Niederrer and Schecker:[57]

> *"Operational descriptions give reports of student behavior in specific situations. Theoretical descriptions try to make inferences about conceptions, concepts, ideas, schemes, or abilities students show over a range of different situations." – Niedderer & Schecker (1992)*

Confusing related quantities: An investigation of student understanding of balancing provides an example of confusion between related quantities.[58] In the case of extended objects, we found that students explicitly associated the balancing condition with equal forces (or masses, or weights) to either side of a fulcrum; in the case of discrete mass distributions, they properly took into account the location of the forces (or masses or weights) as well as the amounts. The results were not consistent with a single mistaken belief applied across multiple contexts; they were instead indicative of a failure to differentiate related quantities. Specifically, when two familiar concepts, say mass and distance, are linked to synthesize a new concept, it is necessary to recognize when the new concept is relevant and how it differs from the base concepts. In some cases, it may be "obvious" that the synthetic concept is appropriate, in others there may be some ambiguity and defaulting to a more familiar one may occur. In such cases, it is not accurate to infer a mistaken belief that the two concepts (force and torque, mass and volume, velocity and acceleration…) are the same. In fact, learners will frequently produce formal definitions and even apply them in quantitative situations, without reliably choosing the correct concept in all situations.

Failing to recognize implications of choice of system of interest: In a study of student understanding of work and the work-kinetic energy theorem, we found that students

often did not recognize that their analysis depended on the choice of a system of interest, a choice that determines which interactions are considered external (and capable of changing the total energy of the system) and which are considered internal (and capable only of changing the form that energy takes).[59] However, if the consequences are not obvious, warnings to select a system and to maintain that choice throughout an analysis are not likely to be heeded. In this case, we cannot consider that students have firm views about systems that we need to change; rather that their views are ill-formed and sporadically invoked. (It should be noted that, in general, saying that students are *failing to do "X"* does not mean *they are doing "X" but failing in the attempt.* It simply means that they are *not doing* "*X,"* when *"X"* would be appropriate*.)*

Relying on memorized results: In a study of superposition and reflection, we found that students tended to answer questions about pulses reflected from fixed and free ends of strings with memorized rules based on their observations, even though their observations had not spanned a significant enough range of cases for rules to be inferred.[60] Again, it did not appear to be the case that students had a firm model for the phenomenon under study, rather, they mistook observations as the end point in themselves, not the basis for developing and validating a model.

Lacking a verbal interpretation of a relationship: In a study of entropy and the 2nd law of thermodynamics, we found that students were unable to articulate the meaning of the relationship between a change in entropy and the quantity dQ/T.[61] In particular they could not explain how the quantity dQ/T represents an infinitesimal change in entropy accompanying a heat transfer. We concluded that the symbolic relationship was essentially meaningless as a tool for reasoning. (It may be that there is an interpretative or instructional role here for *symbolic forms* – that is, that students could be guided to see the original relationship as analogous to others with the same symbolic form.[62])

Assuming (inverse) proportionality: In a series of studies of work and impulse, students were asked to compare the momenta of two non-identical objects that were subject to the same force over the same distance.[63,64] Rather than recognizing they could use the concept of impulse, many students erroneously concluded that the momenta must be same because the object with greater mass travels with lower speed while the object with smaller mass travels with greater speed. The assumption that the products are the same is a form of *compensation reasoning*. It is (presumably) not the case that students believe these products must always be the same, just that under the circumstances, and in the absence of a compelling alternative, it seems a reasonable assumption.

Favoring specific cases over broad principles: In a study of student understanding of the first law of thermodynamics in the context of ideal gases, we found that many students preferred the relationship $PV = nRT$ to the first law, despite the fact that it was often inadequate for the tasks at hand.[65] It appeared to us that the ideal gas law was considered at least as general and powerful as the first law. This tendency is similar, but not identical, to the tendency to focus on surface features rather than broad principles.[66]

Overgeneralizing: In a study of electric potential difference we asked students about the potential of a grounded conducting sphere that had an induced positive charge in the presence of a nearby negatively charged object.[67] Despite the fact that the sphere was explicitly noted as grounded both in text and graphically, the majority of students failed to recognize that the sphere is at the same potential as ground – namely zero, by definition. Mostly they assumed that since the sphere is positively charged it must be at a positive potential, some even citing the relationship $V = kQ/R$ in support. In other cases, this relationship was invoked to explain the potential of non-spherical objects with non-uniform charge distributions. Thus, students generally failed to recognize its limitation. More importantly, they had not developed a model for the concept of ground, nor a sufficiently general understanding of potential and potential difference to answer questions when a simple formula was inadequate, or inappropriate.

Blending or hybridizing models: In studies of optics, we observed students reasoning about diffraction and interference phenomena with elements of both ray and wave models.[34] For instance, some students attributed a single-slit diffraction pattern to two effects: the broad central maximum was explained as resulting from light traveling in straight lines through the slit while the fringes were explained as resulting from sort of the spreading of light. As one student said during an interview: ''some of the light will go through without being diffracted and then the light that strikes the edges will be diffracted off.'' In the paper that reported these findings, students were said to express "firm beliefs." It was indeed the case that students articulated their thoughts in ways suggestive of having constructed these accounts prior to the interviews, perhaps in a resourceful effort to reconcile the two models they had been taught. Moreover, in response to probing, they tended to provide additional details rather than to express hesitation or doubt. Thus, while presumably not rooted in everyday experience, these ideas seem to exhibit "misconceptions-like" characteristics.

Invoking mistaken beliefs: Mistaken beliefs have been cited in a number of other contexts. In early studies of electric circuits, Shaffer and McDermott documented the mistaken belief that current is "used up" as it traverses the resistive elements in a circuit.[68] This belief was manifested as the tendency to assume the bulb brightness in a

series circuit would vary with increasing distance from one end of battery (usually decreasing with distance from the positive terminal). However, it was clear that students would often, even typically, state that current is conserved and be able to express Kirchoff's laws, but they acted as if they believed that current is used up. This investigation also demonstrated that students did not have a clear understanding of current itself, so it should not be inferred that students necessarily associated the term "current" with a robust alternative model for charge flow. Thus, even in the most "misconceptions-like" cases, there is evidence of a lack of coherence.

5.4   Addressing difficulties

As mentioned above, the *Tutorials* aim to help students construct coherent conceptual frameworks. At the same time they *address specific difficulties*. As should be clear by now, there are many types of difficulty and therefore there is no one-size-fits-all strategy. In particular, "addressing" does not necessarily mean "destroying" – in some cases it *could* mean that, but also, depending on the nature of the difficulty in question, "refining," "redirecting," "broadening," "further specifying," "differentiating," "generalizing," etc. Moreover, each strategy for addressing a specific difficulty is embedded in a broader structure of experiments and exercises intended to guide the construction of a *coherent conceptual framework.* For example, responses to the coin toss problem could be interpreted as evidence of a belief that there is always a force in the direction of motion, which might call for vigorous rebuttal. Alternatively, it might be interpreted as the use of different *p-prims* for different instants during the motion. In either case, the goal is not to have students learn that the net force on a tossed coin is non-zero throughout its path, rather it is to recognize non-zero acceleration in general, to distinguish between a single force and the net force, to properly associate net force with acceleration, etc. In this and most other cases, incorrect responses are symptoms of a larger underlying problem – the lack of a coherent conceptual framework. While some progress is possible by treating various difficulties individually, we have found that a holistic approach is most effective.

An instructional strategy used for some, but not all, difficulties is referred to as *elicit-confront-resolve* (ECR). In the *elicit* phase, a task or question is used that research shows will tend to elicit an incorrect response, which is itself based on incorrect premises or faulty reasoning. For instance, students may be asked to make a prediction about the relative brightness of two lightbulbs connected in series. If they tend to think that current is used up in a circuit, they may predict that "downstream" bulbs are dimmer than those "upstream." In the *confront* phase, students are led to acknowledge an inconsistency between the prediction and an observation or between two logically incompatible conclusions. For example, a simple circuit may be set up and students

asked to record their observations about the lightbulbs. The confront stage is often challenging, as many learners will attempt to find ways around the inconsistency rather than facing it head on. For instance, they may argue that there are slight differences in brightness that validate their predictions. The *resolve* stage is the most critical. Students are guided in identifying the flawed reasoning or incorrect premises that led to the incorrect prediction and revise their ideas in the light of new information. In short, they troubleshoot their own thinking. Absent appropriate support, students may fail to appreciate the flaws in their thinking, or focus only on the situation at hand, not the broader principles involved.

The strategy is not always as direct as in the above example. In the case of balancing, we might have asked students to predict how the amount of mass to the left of the balance point of an extended object compares to that to the right, thereby eliciting the idea that the masses must be equal.[58] We might then have *shown* students that the masses are not equal. (In fact, this was attempted as part of a lecture demonstration with only modest success.) However, we had concluded that there was a need to distinguish torque from force [v]. Therefore, in the tutorial, students suspend an asymmetric, extended object from a point near its center of mass and observe that it rotates away from its initial horizontal orientation. They predictably infer that the masses to the left and right of the suspension point are unequal. They are then asked to restore balance by adding a piece of clay somewhere on the object. Upon doing so, they are again asked about the masses; many assert that now that the clay has been added, the masses are once again equal. They are then asked both to move the clay a short distance and to add (or remove) clay to restore balance. After these experiments they generally agree that it is possible to balance the object with unequal amounts of mass, provided they are carefully arranged. Thus, they are led to recognize that there are two variables that must be taken into account: the amount of mass, and its location relative to the fulcrum.

In the physics education community, cognitive conflict as an instructional strategy is tightly linked to the notion of misconceptions. Thus ECR has been viewed by some as an effort to help students "accommodate" new conceptions in the ways articulated by Posner *et. al.*[44] In short, ECR has been viewed as reflecting an assumption that students hold robust alternative beliefs that are incommensurate with scientific beliefs and that must be discarded (if not destroyed) for progress to be made. Leading students to make errors and exposing those errors are indeed key parts of the ECR strategy. However, I argue that it is not necessary to assume that those errors stem from pre-existing incorrect beliefs. Nor does resolving a conflict necessarily imply replacing a faulty

---

[v] Here we use the terms mass, weight and force interchangeably as the difference is not germane to the issue at hand, and to do otherwise would be unnecessarily cumbersome.

conceptual system with a more canonical one. Generally speaking, in PER cognitive conflict has been justified by theories of robust alternative conceptions, but the theoretical grounds for this link are weak. For instance, Smith, diSessa, and Roschelle pointed out that, if prior knowledge is demolished, then it is difficult to simultaneously claim that it is the foundation for the construction of new knowledge.[49]

Even if one concludes that *in some cases* learners have developed firm but mistaken beliefs, it does not follow that they must be made to acknowledge those beliefs and be presented with evidence that is incompatible with those beliefs. Instead, we might hypothesize that it would be better to ignore those beliefs because explicit attention runs the risk of reinforcing them via the *familiarity effect*. Moreover, reasoning about physics situations might *always* involve a competition between previously constructed ideas and those presented in formal instruction – in which case the latter could be reinforced and made more attractive in the hopes that they might take precedence over the mistaken beliefs, rather than replacing them. Likewise, inducing cognitive conflict might be seen as a valid strategy for promoting the development of strategies for conflict detection and resolution, and not just to convince learners that their initial ideas are unproductive. Therefore, while it may well be the case that ECR was originally inspired by theories of conceptual change, its continued use does not imply adherence to such theories.

It should also be recognized that ECR has been criticized on other grounds. One point of criticism is that repeated incorrect predictions or answers will have a corrosive effect on students' self-efficacy, interest and motivation. While this would obviously be a source of concern, to my knowledge there is no evidence that this is a widespread or lingering effect of using tutorials in a physics course. In fact, courses with tutorials often do better on surveys of attitudes and expectations than traditionally taught courses.[69] That is not to say that there are widespread *improvements,* only that the trends are somewhat less negative than is typically observed. To be sure some students express dismay about their predictions being wrong. Others appreciate that these errors occur in low-stakes situations. My own experience as a learner (and as a researcher, teacher and curriculum developer) has led me to believe that a certain degree of frustration and confusion is an inevitable part of trying to grasp something difficult. Learning how to test one's own understanding, how to find the flaws in one's own arguments, and how to carefully delineate different lines of argument that are in conflict is an essential element of critical thinking. Thus it seems to me that eliciting students' ideas could promote their ability to reflect on their own thinking, and perhaps move in the direction of monitoring that thinking on their own; in short, to be more "metacognitive." However, the extent to which this happens, if at all, is unknown. In any case, learning does not proceed smoothly forward any more than science does. We

need students to be as attuned to the unexpected in the classroom as we are in research. I believe it is possible to foster an environment in which incorrect predictions are viewed not as failures of intuition but as signals of a breakthrough to come. However, I recognize that it is not obvious how we can support students through this process.

A related criticism of ECR has to with its potential to harm students' emerging grasp of the epistemology of physics. Specifically, it has been suggested that the experience of finding their commonsense ideas to be at odds with, or at least inadequate for, formal physics will counteract any efforts we make to convince students that formal physics is continuous with everyday thinking. This may well be true. Evidence one way or another would be interesting. Studies of experts indicate that intuition plays a critical role in thinking about new situations.[70] To my mind, helping students to recognize the dual roles of intuition and formal reasoning is the next big challenge in introductory physics. To start with, we need to help students understand the difference, as some of my colleagues and I are attempting.[71-73]

Finally, another source of criticism relates to the demand that students make predictions that we know they will probably generate on the basis of insufficient or incorrect grounds, whereas some argue that scientists make predictions on the basis of careful analysis and well-justified hypotheses, if not fully developed theories. This is not entirely true, as a certain degree of playful undirected exploration occurs in many scientific settings. Nevertheless, we do not generally ask students to do anything that they cannot in principle do successfully. Rarely, if ever, are they asked to make predictions that would be impossible on the basis of what they have been taught or can expected to sensibly assume. Thus, the predictions that often feature in ECR are not guessing, they are tests of the current state of students' knowledge.

6. "Instructional significance"

At the outset of the tutorial project, we would frequently assess the impact of an individual tutorial by asking students to answer a "post-test" question that they hadn't seen before, but that tested their ability to reason qualitatively with the concepts targeted in that tutorial. If they (as a group) performed better than they (or a similar group) had performed on a related "pretest" prior to the tutorial, we judged their (average) level of understanding to have improved. The definition of understanding implied by this practice is that:[53]

> *"…a student will be considered to understand a topic if, when faced with an unfamiliar problem, he or she reliably selects the appropriate concepts and principles, applies them correctly, and constructs a logically sound*

> *solution. Accordingly, instruction will be considered to be improved if more students do this more frequently. – Heron, 2003*

This approach reflects several assumptions. Chief among these is that a process called *learning* results in something called *understanding,* which can in turn be measured by performance on a test. It is certain that there are forms of learning that lead to changes not detectable by current test practices, no matter how thoughtful or theoretically justified. I have observed students arguing in the classroom, grappling with significant conceptual issues and ultimately reaching well-justified conclusions. It does not follow that each and every one of the participating students will subsequently correctly apply the concepts in question on a test. Nevertheless, I think we are justified in using improved performance as *a* measure of improved instruction, especially if we reserve judgment about individual students and take a statistical approach. Specifically, I believe we can claim that the level of understanding in a group is higher on average if more members of the group are able to answer a question correctly. To be clear, a student who answers a question correctly does not necessarily understand better than one who does not, just that a group in which more students answer correctly has a higher average level of understanding than a group in which a smaller proportion does. It must be emphasized that experience tells us something about the time scales over which advances stabilize, but the underlying processes are not well understood. Moreover, it is clear that learning does not always advance in the "forward" direction – some steps that might appear backwards to the expert may be natural parts of the process of constructing understanding.[74] Therefore care must be taken that judgments of performance separated by time are not too close for any meaningful gains to have been consolidated.

The approach outlined above depends on comparing the performance of two different groups of students on two different sets of tasks in two different instructional circumstances (with and without an intervention). An alternative is to use the *same* tasks with two groups randomly assigned to either participate in the intervention or not. In principle, any differences between the groups that might affect either their starting point or their responsiveness to instruction will be controlled. In so-called "ideal" conditions, the time on task will also be controlled such that one group has an equivalent amount of attention to the topics under investigation but perhaps in a conventional way. However, this scheme raises ethical issues in real classrooms – how do we justify randomizing students to different conditions that may affect their grades and, possibly, preparation for subsequent courses? It seems to me preferable to offer the most fully justified instruction to *all* students and figure out how to make meaningful measurements under those circumstances. Moreover, once classes are assigned, it is essentially impossible to ensure that all subsequent conditions are the

same, save for the one under examination. However, for a valid experiment, it is not necessary to control variables for which there are no theoretical or empirical grounds for expecting a significant effect. It has been our experience that many (most, in fact) of the variables that one might consider controlling can be safely ignored, provided we restrict our attention to **instructionally significant** differences in performance. Such differences are large enough to dominate those of other, lower-impact variables, and large enough to justify changes to instructional practice. For example, we have found it to be reasonable to compare different groups at different stages of instruction, even if the two groups have different instructors, meet at different times of day, use different textbooks, etc.*,* provided (1) the student population is the same (*e.g.,* both introductory calculus-based classes at UW) and (2) only differences larger than about 10% are considered instructionally significant.[75]

The first approach outlined earlier also reflects the implicit assumption that responses to related but different tasks can be compared directly. While we endeavor to ensure that post-test questions are more difficult than pretests, we have not traditionally considered the role that reasoning plays in determining whether a question will be easy or difficult, regardless of the putative content. For instance, it is not necessarily the case that the reasoning required to *infer* how a circuit of batteries and bulbs is connected, given information about relative bulb brightness, is equally challenging as the reasoning required to *predict* the relative brightness of bulbs, given their electrical connections. (I will return to this issue below.)

It may be obvious that tasks that seem equivalent to the expert may elicit different responses, but the alternative has often been to use the *same* questions both before and after an instructional intervention (as is still commonly done with the FCI and other multiple-choice tests). It has been argued that students merely learn the answers and that no underlying change in understanding has necessarily taken place. This is especially true for multiple choice questions; less so if judgments are made about the quality of the explanations provided. It is also less true if an entire battery of questions is asked rather than one or two. The former is typical in evaluating course-level changes; the latter more typical in evaluating a single, targeted intervention. In either case, repeating the same measures is most justified when there is no reason to expect that the mere fact of repetition will influence the results. Therefore, I maintain that it is optimal if students do not see the exact same task twice.

The system that has evolved at UW involves *repeatedly* using the same tasks with different groups of students under different instructional circumstances: the tasks are given before a given tutorial in several classes and after the tutorial in several others. Having multiple instances of each condition allows the impact of uninteresting

variables to be gauged and interesting effects to be clearly seen. This scheme is imperfect but uses both scientific and statistical methods to establish the trustworthiness of the conclusions.

Some of the pre-test/post-test schemes outlined above invite interpretation in terms of a binary view of student understanding. Specifically, testing students with different questions before and after an intervention is consistent with the following beliefs: (1) students will apply laws, rules and definitions in a logical fashion to proceed from given information to valid inferences and (2) once a misconception has been eradicated, it is gone, replaced by a new, better conception that provides the initial premises for that logical reasoning. I don't believe that my colleagues viewed student learning in such simplistic terms. Certainly, their writing at the time the tutorials were initiated suggests otherwise. However, these practices are surely responsible for creating, at least in part, the perception that the UWPEG subscribes to a misconceptions view of student thinking.

7. Current directions

Before I joined the field of PER as a postdoc, I served frequently as a graduate TA in introductory courses. Each week I was expected to teach groups of freshman students how to apply some sort of formalism to analyze a classic situation: e.g., applying energy conservation to a modified Atwood's machine, Kirchoff's laws to a two-loop circuit, Gauss's law to a spherical charge distribution. In Canada, we called these sessions "tutorials" but the closest analogy in the American system is a recitation. At a certain point during each session students would declare themselves ready for a quiz. Later, when I graded those quizzes, I was baffled by their performance. Then, as now, the central questions that keep me up at night are "Why do students make errors? Specifically, why do they often answer questions in ways that directly contradict what they have been taught? How do learners who initially apply ideas incorrectly, indiscriminately, or non-rationally become competent?"

When I was nearing the completion of my PhD, Lillian McDermott gave a colloquium in my department. (That night I told my parents that I knew what I was going to do with my life.) The notion that one could systematically investigate student learning was a revelation. I soon learned about misconceptions. While the notion of firm alternative beliefs resistant to instruction offered a very welcome explanation for my experiences as a TA, from my earliest days at UW I was faced with abundant evidence of the inadequacy of this view. Working side-by-side with Lillian convinced me that the notion that students have developed sound (within limits) frameworks in which well-defined concepts are merely arranged incorrectly is simply not plausible. It became

important to me to resist the temptation to attribute errors to mistaken beliefs that exist within students' minds, waiting only to be called forth when needed to answer a physics question. It might be the case that after a few decades of experience of pushes, pulls and moving objects, students would have concluded that a force is needed to maintain motion. But where would incorrect ideas about capacitance, electric potential difference, the photoelectric effect or matter waves come from? The sections above outline my current thinking about these issues. But there are other aspects of the search for the causes that underlie common errors that are at the forefront of my thinking at the moment.

To begin, it is essential to recognize that not all errors arise in the same way, even those that seem superficially similar. In some cases, students seem to use an incorrect premise in a logical way to arrive at an answer. For example, when asked to compare the brightness of bulbs connected in a series circuit, some students will state that brightness is an indication of current, that current is used up as it flows through a circuit, and therefore each subsequent bulb in a series arrangement will be dimmer than the one before it. In this case an incorrect premise (current is used up) is applied logically – the conclusion is sound, given that premise. In other cases, the logical chain that relates the answer to any supporting statements made is unclear. For example, when asked to predict where blocks of identical volume but different mass will come to rest in a tank of water, some students will claim that blocks will rest, completely submerged but above the bottom of the tank, because "the buoyant force is greater at greater depth." In this case, the statement is incorrect; but the significance of it in explaining the prediction is not spelled out completely, nor is it easy to infer. It may be the case that the statement and the prediction are only loosely related and it would be a mistake to assume that one led to the other. This tendency – to infer a straightforward thinking process from the incomplete and often fragmentary and incoherent writings and utterances of students – is tempting but dangerous for researchers. Even for the same response, students may arrive at their answers in very different ways. It is also clear that different reasoning processes that may lead students to answer incorrectly even when they might have answered correctly under slightly different circumstances. It was not until recently that I became aware that dual-process theories may provide an explanation.[76]

My work with Mila Kryjevskaia, MacKenzie Stetzer, Beth Lindsey and Andrew Boudreaux attempts (in part) to understand how students can use correct premises and logical reasoning – meaning the methodical drawing of valid inferences – to support correct answers to one part of a multi-part question, but provide patently inconsistent and poorly justified answers to the next. Mila and Mac have demonstrated that the results make sense if we accept that in some cases, a student quickly generates an

intuitive response and then concocts an explanation to rationalize that response.[73] Thus it is not necessarily the case that the answer is the product of logical reasoning, at least not the reasoning that is offered in its defense. Understanding when reasoning is employed, and when intuition is instead judged sufficient, is my current preoccupation. As noted earlier, I believe that we can make significant progress if we can help students to become aware of the dual roles of reason and intuition in learning (and, in fact, creating) physics.

The degree to which framing – or the expectations that accompany a physics task – affects responses is an increasingly important and related issue.[51] When students enter into an interview, take a test, or work with a group in class, their expectations about what sort of thinking and behavior are called for influence that thinking and behavior. The reasoning that takes place cannot be viewed in isolation from the circumstances that called it forth. Given that we have no desire to remove context – to make classroom learning even more disjoint from the real world, we must do what we can to understand and leverage these fundamental facts of learning.

I am aware that some people think a focus on errors is dangerous. One line of argument points out the risk of missing the competence that students exhibit even when they are producing incorrect answers. This is a reasonable concern. However, the tutorials intentionally build on students' strengths. Even if they have not grasped the meaning of lectures or readings, tutorials assume student competence in logical deductive and (occasionally) inductive reasoning. If a goal of instruction is to help students reason correctly with the ideas covered in the curriculum, then a focus on those parts where they fall short is natural.

I also argue that errors play a critical role even in the resources framework. Thus far, at least, resources have chiefly been identified in responses to questions that require qualitative reasoning and that are known to have high error rates. In fact I am engaged in just such an investigation with Amy Roberston, Rachel Scherr and Lisa Goodhew.[77] Therefore I argue that a focus on errors does not treat students as incapable or inferior. I recognize that speaking about what "learners" or "students" do or think positions me in judgment of their efforts and accomplishments. For what it is worth, everything I have said applies to my own experiences as a learner as well. I assume that students have reasons for responding the way they do, only that their reasons might not withstand much scrutiny. Meeting students where they are entails a careful exploration of their thinking, noting both the weaknesses and the strengths.

8. Conclusion

It may seem impossible to develop instructional materials without clear and precise theoretical guidance, but the alternative is not to proceed by trial and error. Such a process would be grossly inefficient and hard to justify ethically. Yet even the most developed theories don't prescribe what steps to take in specific circumstances. As Burkhardt and Schoenfeld stated[78],

> *"General theories are weak, providing only general guidance for design; nonetheless they receive the lion's share of attention in the research literature. By overestimating theory's strength… damage has been done. … In the case of constructivism, a naïve faith in 'cooperative learning' as a mechanism for generating productive conversations has sometimes resulted in pleasant but content-free discussions." - Burkhardt & Schoenfeld, 2003*

I think we, as a field, have progressed since this point, as experimental precision and theoretical specificity move towards convergence. I am optimistic about recent advances in bridging the divide between the "frictionless experiments" of cognitive science (in Joe Redish's words) and classroom teaching.[79] Nevertheless, as far as the development of instructional materials goes, the UWPEG continues to draw from theory where it seems most robust, to draw on the experience of perceptive teachers, and to exploit regularities that occur across topic areas, even without a clear understanding of underlying mechanisms.

To conclude, I think it is useful to return to the pragmatic position. Some words from diSessa are relevant here:[80]

> *I think that the very simple idea of engaging students' ideas is surprisingly powerful, independent of particular theoretical direction. So, many of the good properties of conceptual change inspired instruction come independent of specific theories of students' ideas. … many instructional methods are sensible within both theoretical perspectives [theory theory and knowledge in pieces], if for different reasons. – diSessa, 2008*

I would add that the accumulated expertise of PER cannot necessarily be distilled into well-defined principles. As much as theory can inform research and curriculum development, so can "craft knowledge" and "clinical wisdom." As in all scientific pursuits, creativity and insight do not follow from well-defined procedures. Nevertheless, I have grappled seriously with the challenges of building, interpreting

and applying theory. I have read widely and engaged with people in the US-based and international PER communities. I have taken challenges seriously, been pressed to reflect on the nature of "difficulties," and have reflected on the breadth of strategies in *Physics by Inquiry* and *Tutorials in Introductory Physics.* While my views have certainly evolved as a result, I continue to take a cognitivist view of learning and remain fundamentally interested in how people develop the concepts and principles of physics. Increasingly I am interested in how their reasoning strategies either support or work against these efforts. Finally, on (much) reflection, I remain convinced that "identifying and addressing difficulties" is a practical, and ethical, approach to improving student learning.

Afterword

This article was written in response to a request to describe how I "conceptualize PER." In my view, "conceptualizing" involves defining the prevailing concepts, describing how they are related to one another, and perhaps proposing an underlying mechanism for those relationships in terms of entities whose properties lead to observable phenomena, even while they themselves remain unseen. In terms of PER I think this would mean to explain how the field itself operates – to answer the questions: Who does PER? What do they study, and to what ends? How do they study it, and with whom? Why do some ideas and findings prove influential while others do not? What external forces (societal changes, technological advances, political movements) drive the trends? What unseen forces drive those same trends in ways the actors may not perceive? This charge proved too ambitious for me. Instead I chose to focus on what I know best – my own corner of the field. I am not sure I have succeeded in explicating a coherent conceptual framework for even that limited purview as I detect contradictions in what I have written and in a few places my memory proved faulty. But I have tried to define some concepts and principles that drive my own work, and to some extent that of my colleagues, although their perceptions might differ. I trust that the result will be of interest to others in the field, and that it may stimulate reflection, especially among current and former UWPEG members, whose recollections and interpretations surely differ and whose views I look forward to hearing.

Acknowledgments

References

[1] McDermott, L. C., Shaffer, P. S. & The PEG at UW, *Tutorials in Introductory Physics*. (Prentice Hall, 2002).

[2] Meltzer, D. E. & Otero, V. K, "A brief history of physics education in the United States," *Am. J. Phys.* **83**, 447 (2015).

[3] McDermott, L. C, "Melba Newell Phillips Medal Lecture 2013: Discipline-Based Education Research— A View From Physics," *Am. J. Phys.* **82,** 729–741 (2014).

[4] Cummings, K. in *The National Academies' Board on Science Education Commissioned Papers* (2011).

[5] Hake, R. R, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.* **66**, 64(1998). doi:10.1119/1.18809

[6] Mazur, E, *Peer Instruction* (Pearson, 1997).

[7] Laws, P. W, *Workshop Physics Activity Guide* (John Wiley & Sons, Inc., 1995).

[8] Sokoloff, D. R. & Thornton, R. K, *Interactive lecture demonstrations : active learning in introductory physics*. (John Wiley & Sons, 2004).

[9] Sokoloff, D. R., Thornton, R. K. & Laws, P. W, *RealTime physics. Module 1, Mechanics : active learning laboratories*. (Wiley, 2012).

[10] Van Heuvelen, A, *ALPS Kit for Mechanics*. (Hayden-McNeil Publishing, 1996).

[11] Heller, P. & Hollabaugh, M, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups," *Am. J. Phys.* **60,** 637–644 (1992).

[12] Maloney, D. in *Handbook of research on science teaching and learning* (ed. Gabel, D. L.) 327–356 (Macmillan, 1993).

[13] Hestenes, D., Wells, M. & Swackhamer, G, "Force concept inventory," *Phys. Teach.* **30,** 141–158 (1992).

[14] McCullough, L. & Meltzer, D. E, "Differences in Male/Female Response Patterns on Alternative-format Versions of the Force Concept Inventory," in *2001 Physics*

Education Research Conference Proceedings (American Association of Physics Teachers, 2001). doi:10.1119/perc.2001.pr.013

[15] Redish, E. F., Saul, J. M. & Steinberg, R. N, "Student expectations in introductory physics," *Am. J. Phys.* **66**, 212 (1998). doi:10.1119/1.18847

[16] Redish, L. C. M. F. & Mcdermott, L. C, "Resource Letter : PER-1 : Physics Education Research," *Am. J. Phys.* **67,** 755 (1999).

[17] *Proceedings of the 1998 Physics Education Research Conference, University of Nebraska-Lincoln, August 1-2*.

[18] *Proceedings of the 2003 Physics Education Research Conference* (American Institute of Physics, 2004).

[19] Engelhardt, P. V., Churukian, A. D. & Rebello, N. S, *2012 Physics Education Research Conference (Philadelphia, PA, August 1-2, 2012)* (American Institute of Physics, 2013).

[20] Brewe, E., Kramer, L. & Sawtelle, V, "Investigating student communities with network analysis of interactions in a physics learning center." doi:10.1103/PhysRevSTPER.8.010101

[21] Otero, V. K. & Harlow, D. B. in *Getting Started in PER, Vol. 2* (2009).

[22] Heron, P. R. L, " 'Thinking like a physicist' about physics education." (2017).

[23] McDermott, L. C, "Combined Physics Course for Future Elementary and Secondary School Teachers," *Am. J. Phys.* **42,** 668–676 (1974).

[24] McDermott, L. C, "Teacher education and the implementation of elementary science curricula," *Am. J. Phys.* **44,** 434–441 (1976).

[25] McDermott, L. C., Rosenquist, M. L. & van Zee, E. H, "Strategies to improve the performance of minority students in the sciences," *New Dir. Teach. Learn.* **1983,** 59–72 (1983).

[26] McDermott, L. C., Shaffer, P. S. & The PEG at UW, *Tutorials in Introductory Physics*. (Prentice Hall, 1999).

[27] Articles, dissertations and talks documenting the impact of Tutorials in Introductory Physics. Available at: https://courses.washington.edu/uwpeg/sites/default/files/files/V2 - TOC.pdf.

[28] Articles and dissertations motivated by Tutorials in Introductory Physics. Available at: https://courses.washington.edu/uwpeg/sites/default/files/files/V4 - TOC.pdf.

[29] McDermott, L. C. & The PEG at UW, *Physics by Inquiry*. (John Wiley & Sons, Inc., 1996).

[30] Sherwood, B. A, "Pseudowork and real work," *Am. J. Phys.* **51,** 597–602 (1983).

[31] Arons, A. B. "Development of energy concepts in introductory physics courses," *Am. J. Phys.* 1063–1067 (1999).

[32] Jewett, J. W, "Energy and the Confused Student I: Work," *Phys. Teach.* **46,** 38–43 (2008).

[33] McDermott, L. C, "Guest Comment: How we teach and how students learn—A

mismatch?" *Am. J. Phys.* **61,** 295–298 (1993).

[34] Ambrose, B. S., Shaffer, P. S., Steinberg, R. N. & McDermott, L. C, "An investigation of student understanding of single-slit diffraction and double-slit interference," *Am. J. Phys.* **67,** 146–155 (1999).

[35] Wosilait, K., Heron, P. R. L., Shaffer, P. S. & McDermott, L. C, "Addressing student difficulties in applying a wave model to the interference and diffraction of light," *Am. J. Phys.* **67,** S5–S15 (1999).

[36] Close, H. G. & Heron, P. R. L, "Student understanding of the angular momentum of classical particles," *Am. J. Phys.* **79,** 1068–1078 (2011).

[37] Shaffer, P. S. & McDermott, L. C, "A research-based approach to improving student understanding of the vector nature of kinematical concepts," *Am. J. Phys.* **73,** 921–931 (2005).

[38] McDermott, L. C, "Millikan Lecture 1990: What we teach and what is learned— Closing the gap," *Am. J. Phys.* **59,** 301–315 (1991).

[39] Halloun, I. A. & Hestenes, D, "Common sense concepts about motion," *Am. J. Phys.* **53,** 1056–1065 (1985).

[40] Disessa, A. A, "Unlearning Aristotelian Physics: A Study of Knowledge-Based Learning," *Cog. Sci.* **6,** 37–75 (1982).

[41] Clement, J, "Students' preconceptions in introductory mechanics," *Am. J. Phys.* **50,** 66–71 (1982).

[42] Ogborn, J, "Constructivist Metaphors of Learning Science," *Sci. Educ.* **6,** 121–133 (1997).

[43] Driver, R. *et al.* "Constructing Scientific Knowledge in the Classroom," *Educ. Res.* **23,** 5–12 (1994).

[44] Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A, "Accommodation of a scientific conception: Toward a theory of conceptual change," *Sci. Educ.* **66,** 211–227 (1982).

[45] Thornton, R. K. & Sokoloff, D. R, "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula" *Am. J. Phys* **66**, 338 (1998). doi:10.1119/1.18863

[46] Redish, E. F, *Teaching physics : with the physics suite*. (John Wiley & Sons, 2003).

[47] McDermott, L. C, "Research on conceptual understanding in mechanics," *Phys. Today* **37,** 24–32 (1984).

[48] DiSessa, A, "Toward an Epistemology of Physics," *Cogn. Instr.* **10,** 105–225 (1993).

[49] Smith III, J. P., diSessa, A. A. & Roschelle, J, "Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition," *J. Learn. Sci.* **3,** 115–163 (1994).

[50] Hammer, D, "More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research," *Am. J.*

Phys. **64,** 1316–1325 (1996).

[51] Hammer, D. *et al.* in *Transfer of Learning from a Modern Multidisciplinary Perspective.* (ed. Mestre, J.) 89–119 (IAP, 2004).

[52] Driver, R. & Erickson, G, "Theories-in-action: Some theoretical and empirical issues in the study of students' conceptual frameworks in science," *Stud. Sci. Educ.* **10,** 37–60 (1983).

[53] Heron, P. R. L, "Empirical Investigations of Student Understanding," *2003 PERC Proc. [Madison, WI, August 6-7, 2003] AIP Conf. Proc.* **720,** 15–18 (2004).

[54] *Research on Physics Education: Proceedings of the International School of Physics 'Enrico Fermi' Course CLVI.* (IOS, 2003).

[55] Scherr, R. E. & Heron, P. R. L, "Education metaphors we live by," *2016 Physics Education Research Conference Proceedings* 316–319 (American Association of Physics Teachers, 2016). doi:10.1119/perc.2016.pr.074

[56] Robertson, A. D., Goodhew, L., Scherr, R. E. & Heron, P. R. L, "University student conceptual resources for understanding forces," *Physics Education Research Conference Proceedings (Cincinnati, OH, July 26-27, 2017)* (2017).

[57] Niedderer, H. & Schecker, H, "Towards an explicit description of cognitive systems for research in physics learning" in *Research in Physics Learning - Theoretical Issues and Empirical Studies (Proceedings of an International Workshop in Bremen)* (eds. Duit, R., Goldberg, F. & Niedderer, H.) 74–98 (IPN, 1992).

[58] Ortiz, L. G., Heron, P. R. L. & Shaffer, P. S, "Student understanding of static equilibrium: Predicting and accounting for balancing," *Am. J. Phys.* **73,** 545 (2005).

[59] Lindsey, B. A., Heron, P. R. L. & Shaffer, P. S, "Student understanding of energy: Difficulties related to systems," *Am. J. Phys.* **80,** 154–163 (2012).

[60] Kryjevskaia, M., Stetzer, M. R. & Heron, P. R. L, "Student understanding of wave behavior at a boundary: The limiting case of reflection at fixed and free ends," *Am. J. Phys.* **79,** 508–516 (2011).

[61] Cochran, M. J. & Heron, P. R. L, "Development and assessment of research-based tutorials on heat engines and the second law of thermodynamics," *Am. J. Phys.* **74,** 734–741 (2006).

[62] Sherin, B. L, "How Students Understand Physics Equations," *Cogn. Instr.* **19,** 479–541 (2001).

[63] Lawson, R. A. & McDermott, L. C, "Student understanding of the work-energy and impulse-momentum theorems," *Am. J. Phys.* **55,** 811–817 (1987).

[64] Pride, T. O., Vokos, S. & McDermott, L. C, "The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems" *Am. J. Phys.* **66**, 147 (1998). doi:10.1119/1.18836

[65] Loverude, M. E., Kautz, C. H. & Heron, P. R. L, "Student understanding of the first law of thermodynamics: Relating work to the adiabatic compression of an ideal gas," *Am. J. Phys.* **70,** 137–148 (2002).

[66] Chi, M. T. H., Feltovich, P. J. & Glaser, R, "Categorization and representation of physics problems by experts and novices," *Cogn. Sci.* **5,** 121–152 (1981).

[67] Hazelton, R. L. C., Shaffer, P. S. & Heron, P. R. L, "Facilitating model-building of electrostatics concepts related to conductors," *2015 Physics Education Research Conference Proceedings Proceedings (College Park, MD, July 29-30, 2015)* 131–134 (2016). doi:10.1119/perc.2015.pr.028

[68] McDermott, L. C. & Shaffer, P. S, "Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding," *Am. J. Phys.* **60,** 994–1003 (1992).

[69] Finkelstein, N. D. & Pollock, S. J, "Replicating and understanding successful innovations: Implementing tutorials in introductory physics," *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **1,** 10101 (2005).

[70] Clement, J, "Imagistic simulation and physical intuition in expert problem solving," *The Sixteenth Annual Meeting of the Cognitive Science Society* (Lawrence Erlbaum Associates, Inc., 1994).

[71] Heron, P. R. L, "Testing alternative explanations for common responses to conceptual questions: An example in the context of center of mass," *Phys. Rev. - Phys. Educ. Res.* **13,** (2017).

[72] Gette, C. R., Kryjevskaia, M., Stetzer, M. R. & Heron, P. R. L, "Probing student reasoning approaches through the lens of dual-process theories: A case study in buoyancy," *Phys. Rev. - Phys. Educ. Res. (in press).*

[73] Kryjevskaia, M., Stetzer, M. R. & Grosz, N, "Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics," *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **10,** 1–12 (2014).

[74] Siegler, R. S, "U-Shaped Interest in U-Shaped Development-and What It Means," *J. Cogn. Dev.* **5,** 1–10 (2004).

[75] Heron, P. R. L, "Effect of lecture instruction on student performance on qualitative questions," *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **11,** 1–14 (2015).

[76] Kahneman, D, *Thinking, fast and slow*. (Farrar, Straus and Giroux, 2011).

[77] Goodhew, L., Robertson, A. D., Scherr, R. E. & Heron, P. R. L, "Student conceptual resources for understanding mechanical wave propagation," *2017 Physics Education Research Conference Proceedings (Cincinnati, OH, July 26-27, 2017)* doi:10.1119/perc.2017.pr.032.

[78] Burkhardt, H. & Schoenfeld, A. H, "Improving educational research: Toward a more useful, more influential, and better-funded enterprise," *Educ. Res.* **32,** 3–14 (2003).

[79] Heckler, A. F. in *Psychology of Learning and Motivation - Advances in Research and Theory* **55,** 227–267 (2011).

[80] diSessa, A. A. in *International Handbook of Research on Conceptual Change* (Routledge, 2013). doi:10.4324/9780203154472.ch2