WIDER: EAGER: **Recognizing, assessing, and enhancing evidence-based instructional practices in STEM at Arizona State University, Polytechnic**

I. Overview

This two-year project is a systematic effort to recognize, assess, and enhance evidence-based STEM instructional practices on the Arizona State University Polytechnic campus, one of four campuses of Arizona State University (ASU). ASU enrolls over 70,000 students across its campuses, and awards over 13,000 undergraduate and graduate degrees each year. ASU is among the nation's leading post-secondary institutions in number of degrees awarded to Hispanic and Native American students. The lead academic unit will be the College of Technology and Innovation (CTI), which is the college responsible for almost all STEM instruction on the ASU Polytechnic campus. CTI offers courses in physics, chemistry, biology, mathematics, engineering, and technology, with degree programs focused on engineering, technology, and applied sciences. At the present time, there is no institution-wide effort to document and/or track the use of evidence-based instruction in STEM undergraduate programs at ASU, although individual instructors in individual courses have occasionally carried out such efforts. This project will document present practices and lay the basis for enhancing future employment of evidence-based instruction.

This proposal is being submitted under the EAGER program because it involves untested, but potentially transformative, research ideas or approaches. In particular, we must apply tentative and so-far unproven methods of recognizing and assessing evidence-based instructional methods in STEM fields outside of physics, by using characteristics and methods that have so far only been validated *rigorously* for physics instruction. We will also be using diagnostic instruments for assessing student learning that—again, apart from physics—are still in various early stages of testing and validation. This proposal is, therefore, better fit to EAGER than to existing programs.

II. Results from Prior NSF Support

PI David Meltzer has received NSF funding as PI or Co-PI on nine separate education-related projects (11 overall), including two during the past five years. Project titles, award numbers, and key publications are listed in the biography pages. Recent work has focused on development of active-learning, inquiry-based curricular materials for thermodynamics in college-level physics and chemistry courses. He and his collaborators have given more than 50 invited and contributed presentations and published more than a dozen refereed papers related to this work; two doctoral dissertations were supported by these grants. The project team has widely disseminated dozens of research-based curricular worksheets, diagnostic test questions, and other instructional materials. Strong evidence for effectiveness of the materials is contained in the cited references. For example, use of our research-based tutorial on the second law of thermodynamics resulted in score gains on diagnostic exams from ≈ 10% correct without tutorial use, up to ≈ 55% correct with use, with similar results reported at two different universities [Christensen, Meltzer, and Ogilvie, 2009]. Most of the work produced by these projects, including papers, presentations, and curricular materials, may be viewed and downloaded directly; see links in [Thermo PER, 2012].

Co-PI Michelle Zandieh has been PI or Co-PI on three NSF-supported projects; the most recent one is NSF-DRL #0634099, "Collaborative Research: Investigating Issues of the Individual and the Collective along a Continuum Between Informal and Formal Reasoning," $267,119, 08/15/07-08/31/12. This collaboration with Chris Rasmussen focused on mathematical progress as students transition from intuitive to more formal ways of reasoning. In geometry, for example, a framework was developed that structures the role of defining as students progress from informal to more formal ways of reasoning. This offers researchers and instructional designers a structured way to analyze or plan for the role of defining in students' mathematical progress. In other work, using data in differential equations and linear algebra, a framework was developed for making sense of learning and teaching by examining collective classroom activity, disciplinary practices, individual conceptions, and individual participation. In addition, through conducting interviews and watching classroom video data, she and her collaborators analyzed and reported extensively on student thinking about particular mathematical ideas, and used this to create various instructional sequences. The project led to three dissertations, a masters thesis, two Research Experience for Undergraduate awards, and a burgeoning collaboration with colleagues in Mexico. (See Biography pages for eight publications resulting from this project.)

III. Introduction: Defining the Problem

Among the primary challenges in this project are:

(a) *defining* what we mean by "evidence-based instructional practices,"

(b) *recognizing* whether and to what degree such practices are in effect, and

(c) gathering and assessing *evidence and measures* of evidence-based instruction.

We believe that careful analysis of these challenges is crucial to the success of this project, so we will address them in considerable detail. This analysis will provide the basis for the remainder of the proposal.

The "Dear Colleague Letter" contains three key terms: *evidence-based instructional practices*, *evidence-based reforms*, and *effective instructional practice*. These phrases beg many questions. Any STEM instructor, after all, is likely to have some sort of "evidence" regarding the outcome of his or her instruction—students' test scores, homework grades, and the like—and instructors are likely to define "effective instruction" by the degree to which students score well on such measures. These measures and standards are generally (though not always) determined individually by each instructor. In principle, institution-wide reform could result from having all instructors attempt to make individual reforms in their courses and carefully monitor student performance year-to-year on unchanging tests and assignments, demonstrating trends toward improvement. This, however, is *not* the method we will adopt. We believe this method would neglect an enormous body of research and development in STEM education, and would not lead to the steady cumulative progress that characterizes science and technology as a whole.

In the STEM education community a consensus has developed that measures of student learning are needed that are in some way standardized, that transcend the idiosyncrasies of individual instructors, and that take into account the wide variation in preparation of incoming students. Decades of work based on use of such tools have led to significant advances in STEM education. With specific reference to physics, Heron and Meltzer [2005] have noted:

> Researchers have developed instructional materials and methods that have been subjected to repeated testing, evaluation, and redesign. Numerous reports have documented significant and reproducible learning gains from the use of these materials and methods in courses ranging from large-enrollment classes at major public universities to small classes in two-year colleges and high schools.

In this vein, we choose to define *effective instructional practice* as practice that results in substantial student learning as validated by research-based assessment measures as referenced above; *evidence-based instructional practices* as practices whose effectiveness has been validated through such measures of student learning; and *evidence-based reforms* as changes in instructional practices leading to increased use of research-validated methods and materials and/or that can be validated as improving student learning relative to previous or traditional instructional methods.

To be more concrete, we will first focus on the area of physics—the PI's specialty—to outline more detailed responses to these key questions. The primary reference for the following discussion is the PI's recent review paper (co-authored by R. K. Thornton) on Active-Learning Instruction in Physics [Meltzer and Thornton, 2012]. Following the extended discussion on physics, we will outline our plans for carrying out a similar process in mathematics, chemistry, biology, and engineering at the ASU Polytechnic campus.

IV. Project Plan for Physics

A. What are "evidence-based instructional practices" in physics?

Meltzer and Thornton recently published the first comprehensive review of evidence-based instructional practices in physics [Meltzer and Thornton, 2012; see download link]. This 19-page paper is the first-ever systematic survey of research-based physics instruction. It provides: (a) a thorough historical review of the development of active-learning physics instruction in the U.S., (b) a working definition of "research-based active-learning physics instruction," (c) a descriptive list of the key features characterizing this type of instruction, and (d) an annotated listing of 173 primary literature references to evidence-based instructional methods in physics.

Meltzer and Thornton state that research-based active-learning instructional methods in physics (sometimes called "interactive engagement" [Hake 1998]) share the following common features:

> (1) they are explicitly based on research in the learning and teaching of physics;
>
> (2) they incorporate classroom and/or laboratory activities that require all students to express their thinking through speaking, writing, or other actions that go beyond listening and the copying of notes, or execution of prescribed procedures;
>
> (3) they have been tested repeatedly in actual classroom settings and have yielded objective evidence of improved student learning.

B. How can one recognize whether and to what degree evidence-based instructional practices in physics are being used?

Based on analysis of the historical development of, and research on, these instructional methods, as well as examination of the specific methods themselves, Meltzer and Thornton conclude that:

> Research-based active-learning instructional methods in physics…share most or all of the following characteristics:
>
> (a) Instruction is informed and explicitly guided by research regarding students' pre-instruction knowledge state and learning trajectory, including:
>   (1) Specific learning difficulties related to particular physics concepts;
>   (2) Specific ideas and knowledge elements that are potentially productive and useful;
>   (3) Students' beliefs about what they need to do in order to learn;
>   (4) Specific learning behaviors;
>   (5) General reasoning processes.
>
> (b) Specific student ideas are elicited and addressed.
>
> (c) Students are encouraged to "figure things out for themselves."
>
> (d) Students engage in a variety of problem-solving activities during class time.
>
> (e) Students express their reasoning explicitly.
>
> (f) Students often work together in small groups.
>
> (g) Students receive rapid feedback in the course of their investigative or problem-solving activity.

(h) Qualitative reasoning and conceptual thinking are emphasized.
(i) Problems are posed in a wide variety of contexts and representations.
(j) Instruction frequently incorporates use of actual physical systems in problem solving.
(k) Instruction recognizes the need to reflect on one's own problem-solving practice.
(l) Instruction emphasizes linking of concepts into well-organized hierarchical structures.
(m) Instruction integrates both appropriate content (based on knowledge of students' thinking) and appropriate behaviors (requiring active student engagement).

Note: For this proposal, we adopt the above listings of three "common features" (under III) and most or all of the "shared characteristics" (under IV) as our working definition of necessary indicators of evidence-based instructional practices in physics, with suitable modifications for other STEM fields as discussed later in this proposal.

In [Meltzer and Thornton, 2012], the authors provided many detailed examples of each of these characteristics along with relevant references to the research literature. Despite this rather explicit and detailed set of parameters, it is still no easy task to determine, in any given case, whether one has at hand an "evidence-based instructional practice."

In some sense, as noted earlier, any instructional practice that generates some form of "evidence"—be it test scores, observation reports, student or peer evaluations, etc.—can be said to be an evidence-based practice. This is *not* what we take that term to mean, at least in physics. Given the vast research literature developed over the past three decades, the physics education community has reached consensus that *systematic* collection and analysis of quantitative and qualitative student learning data are required to provide reliable evidence of *effective* instructional practices.

Many research-based diagnostic instruments, as well as several observational protocols, have been developed to provide evidence of student learning in physics. The Physics Education Group at the University of Washington has published numerous research-based assessment questions that they have used to probe student learning; baseline data regarding performance on these questions by diverse student populations have been published in the American Journal of Physics [University of Washington, 2012]. A large, annotated collection of diagnostic exams is available [North Carolina State U., 2012]. The most well-known and widely used of these assessments are the Force Concept Inventory (FCI), Force and Motion Conceptual Evaluation (FMCE), Conceptual Survey on Electricity and Magnetism (CSEM) and Brief Electricity and Magnetism Assessment (BEMA). (See [Meltzer and Thornton 2012, Section IV].)

Ideally, of course, one would like to have a definition and guide to "evidence-based instructional practices" so clear that one could simply sit down and observe a class or two and use this guide to determine whether such a practice was in use. In fact, two observation protocols have been developed that attempt to aid a trained observer in making such a determination: The Reformed Teaching Observation Protocol (RTOP) was developed and tested at Arizona State University and has been widely used in STEM education [Sawada et al., 2000; Piburn and Sawada, 2000]. A similar instrument with several additional features is the "UTOP" observation protocol developed at the University of Texas [Walkington et al., 2012; UTOP, 2012]. Both of these instruments have their uses.

However, as Meltzer and Thornton imply, evidence-based instructional practices in physics can be difficult to recognize through straightforward observation because they *necessarily* incorporate specific, topic-dependent research-based instructional materials and methods, that may or may not be recognized by the observer. Quite apart from that, there is no clear quantitative measure of how and in what proportion the various characteristics of effective instruction (enumerated in the list above) need be present in order to make instruction actually effective. To emphasize this point: there is currently no reliable evidence or theoretical basis for claiming that a score of, say, 4 out of 4 on characteristics *e*, *f*, *g*, and *h* on the above list either does or does not outweigh a score of, say, 3 out of 4 on characteristics *a*, *b*, *c*, and *d*. Various instructors may emphasize some characteristics over others, but the fact remains that there is no current research base in physics education that allows any clear resolution of this puzzle. Instead, workers in research-based physics instruction gather and analyze evidence on *specific instructional implementations of specific curricula*.

Certainly it is possible to adopt a different approach. Might it be possible to claim, for example, that "students working together in small groups" is an evidence-based instructional practice? Or, "students engage in a variety of problem-solving activities during class time"? If an observer detects these two characteristics during a class observation, have they confirmed the existence of an "evidence-based instructional practice"? Many might say, "Yes." Our response is that such evidence is suggestive, but absolutely not confirmatory of *effective* evidence-based practices. That is because firm evidence of effective practice always occurs in the context of a large set of closely interlinked characteristics, each characteristic (apparently) tightly dependent on the others for overall instructional success. One can not simply implement characteristics *c*, *d*, *f*, and *j* (for example) and claim possession of evidence-based instruction—not, at least, instruction that is actually effective at improving student learning. Such a claim would be a gross misinterpretation of the research literature. Those who might claim otherwise have the burden of providing evidence to support their claim—and in undergraduate physics education, our comprehensive review in [Meltzer and Thornton, 2012] gives us a strong basis for asserting that such evidence has *not* been published in peer-reviewed journals, popular belief to the contrary notwithstanding.

Based on our extensive review of the literature and 20 years practice in the field, we maintain that at the present time the only reliable claim is that *effective evidence-based physics instruction incorporates one or more of the actual methods and materials that have been validated through published, peer-reviewed research to be, in fact, effective at improving student learning in physics.* **We will adopt this as our operational definition of evidence-based instructional methods in physics.** More than 170 references to such materials are systematically enumerated in [Meltzer and Thornton, 2012]. Apart from or lacking incorporation of such materials, any new methods used would themselves have to be assessed through probes of student learning that then yield positive evidence of improved instruction, as already has been done by the methods described in [Meltzer and Thornton, 2012]. Then these new methods could also be claimed as effective "evidence-based instructional methods." **Methods validated in this manner will also meet our operational definition of evidence-based instructional methods**.

C. How will we gather and assess evidence and measures of evidence-based instruction?

Continuing with our focus on physics, we will follow a four-pronged plan to survey, assess, and enhance the use of evidence-based instruction:

(i) We will collect syllabi and instructional materials for all physics courses offered at ASU Polytechnic and examine them to determine whether and to what degree they incorporate evidence-based instructional practices as outlined in [Meltzer and Thornton, 2012]. This will be done in collaboration with the course instructors who, through discussions with the PI and project team, will help make these determinations. This examination will be supplemented by observations, as described below.

(ii) We will carry out observations of selected classes and laboratories to aid in determination of implementations of evidence-based instructional practices. Video recordings may be used in selected cases. In these efforts we will be aided by use of the UTOP protocol (in preference to RTOP due to UTOP's stronger focus on subject-matter content). However, we will *not* be bound by the UTOP quantitative scoring system because in physics, our analysis suggests that use of research-based instructional materials and methods requires considerably greater weighting than currently allowed for by the UTOP scoring system. Instead of strict quantitative scoring, we will initially attempt to determine whether characteristics of evidence-based instruction, as enumerated in Section IV above, are present in *substantial degree*, *moderate degree*, or *insignificant degree* in each of the targeted courses.

(iii) We will begin regular assessment of student learning in introductory physics courses through pre- and post-instruction use of standard diagnostic instruments such as FCI, FMCE [Thornton and Sokoloff, 1998], CSEM, and BEMA, along with selected items published by the Physics Education Group at the University of Washington. In upper-level courses such as PHY 321 and PHY 331 we will also make use of research-based assessments such as those developed in classical mechanics [Ambrose, 2004] and Electricity and Magnetism [Pollock and Chasteen, 2009]. Creation of a solid baseline of student learning data in all targeted courses will allow us to gauge the effectiveness of future reforms as we increase the use of evidence-based instruction.

(iv) We will work with and support instructors and the department chair to facilitate and enhance the adoption and use of evidence-based instructional materials and practices. We will create a CTI library of evidence-based instructional materials in physics and other STEM fields and acquire texts, workbooks, ancillary materials, multimedia materials, and instructors' guides for a wide and representative range of evidence-based materials for use and reference by CTI faculty and staff. In physics, we have at hand the convenient listing provided in [Meltzer and Thornton, 2012] to guide our acquisitions. We will encourage instructors to consider testing these methods in their own classes and make available to them a wide selection of class-tested materials.

It must be acknowledged that we are not adopting a fixed *quantitative* measure to assess "use" of evidence-based instructional practices. The extended investigation carried out while writing [Meltzer and Thornton, 2012] demonstrated to our satisfaction that no current quantitative evaluation system such as RTOP is sufficiently reliable for our purposes. (This could, of course, change in the future with the development of new evaluation methods.) We believe that qualitative assessments as described in (i) and (ii) above, combined with quantitative assessments of student learning described in (iii), will enable us to facilitate and enhance use of evidence-based methods as discussed in (iv).

D. Plan to enhance use of evidence-based instructional methods in physics

Because our work in physics is further advanced than in the other STEM fields, we are able to incorporate in this proposal additional specific steps for enhancing use of evidence-based instructional methods in physics. We will take this on a course-by-course basis, covering all five primary physics courses taught by CTI (apart from recitations and labs):

1. **PHY 101, Introduction to Physics.** When the PI taught this course he employed research-validated materials from his *Workbook for Introductory Physics* [Meltzer and Manivannan, 2002] as well as materials and methods borrowed from other evidence-based practices. In the future we plan to monitor student learning with assessment questions selected from instruments such as FCI, FMCE, and CSEM, and materials published by the University of Washington.
2. **PHY 111-113, General Physics.** This course currently makes use of the evidence-based online tutor *MasteringPhysics* [Morote and Pritchard, 2009]. We plan to monitor student learning with FCI and/or FMCE, and CSEM and/or BEMA, and promote use of research-based materials such as *Tutorials in Introductory Physics*, *Interactive Lecture Demonstrations* [Sokoloff and Thornton, 1997], and *Student Workbook* for *College Physics: A Strategic Approach* [Knight et al., 2009].
3. **PHY 121, University Physics.** The PI will be teaching PHY 121 in fall 2012 and will be using such evidence-based practices as *Tutorials in Introductory Physics* [McDermott et al., 2002], and

MasteringPhysics. He will be monitoring student learning gains with instruments such as FCI and FMCE. Effectiveness of tutorials is potentially enhanced with use of a teaching assistant (undergraduate or graduate), so funding for such an assistant is included in this proposal.

4. **PHY 321, Vector Mechanics and Vibration**. When the PI taught this course he made heavy use of the research-validated *Intermediate Mechanics Tutorials* [Ambrose 2004; Ambrose and Wittmann, 2012]. The PI regularly administered pre- and post-instruction tests to assess student learning but the data have not yet been analyzed. For this project, we will analyze those data and we will work with the current instructor of that course to facilitate further use and testing of these materials. Class enrollment has grown substantially so it is likely that a teaching assistant will be needed to aid in full implementation. Funding for an assistant is included in this proposal.
5. **PHY 331, Principles of Modern Electromagnetism.** We will work with the course instructor to facilitate use of the Junior-level Electricity and Magnetism course materials developed through research at the University of Colorado [University of Colorado, 2012]. We plan to monitor student learning with BEMA [Ding et al. 2006] and/or the Colorado Upper-Division Electrostatics Assessment [Pollock and Chasteen 2009] or a Gauss's law diagnostic [Singh, 2006].

V. Project Plans in Other STEM Fields

Lacking the detailed road-map available for physics via [Meltzer and Thornton, 2012], our approach in other STEM fields necessarily will be, initially, more ad hoc and eclectic. We will make use of various research-based diagnostic instruments to acquire baseline student learning data, and we will use both UTOP and our list of "common characteristics" derived for physics in Section IV above as a first-approximation guide for assessing evidence-based practices through examination of syllabi and other course materials, as well as classroom observations. Specifically, we will initially attempt to determine whether characteristics of evidence-based instruction, as enumerated in Section IV above, are present in *substantial degree*, *moderate degree*, or *insignificant degree* in each of the targeted courses. We will identify and acquire evidence-based instructional materials in various STEM fields, make them known and available to faculty, and, when appropriate, identify workshops or other professional development opportunities that might help instructors adapt and implement new materials and practices.

A. Mathematics

Co-PI Michelle Zandieh, the CTI Mathematics Program Chair, is a specialist in mathematics education research and will lend significant intellectual support to project efforts in mathematics. A research-based assessment instrument developed with NSF support during the past seven years is the Calculus Concept Inventory (CCI) [Epstein, 2007; Epstein, 2012]. This instrument has been administered to more than 5000 students in 30 states and more than a dozen countries. Zandieh and Co-PI Yun Kang, Assistant Professor, both teach calculus courses and will administer the CCI in their courses beginning in fall 2012. This will begin the process of developing baseline student performance data against which future instructional reforms may be measured. Profs. Zandieh and Kang will carry out analysis of assessment data from these and other mathematics courses, participate in classroom observations, assist in documenting use of evidence-based practices in mathematics courses, and assess future needs for new instructional materials that use evidence-based practices for mathematics courses.

B. Biology

Recent work reported by Haak et al. [2011] showed that student performance in college-level biology courses was significantly improved by consistent use of active-learning instructional methods, analogous to the methods discussed in relation to physics in Section IV above. This gives us confidence to initiate surveys of our biology offerings by probing for characteristics similar to those mentioned in the physics discussion. An example of a new evidence-based approach to college biology is the course being developed by Klymkowsky and collaborators [Klymkowsky, 2012; Klymkowsky and Cooper, 2012]. To develop baseline data regarding student learning, we plan to use the research-based assessment test Biological Concepts Instrument (BCI) [Klymkowsky, Underwood, and Garvin-Doxas, 2010], as well as the Conceptual Inventory of Natural Selection (CINS) [Anderson et al., 2004 (original, 2002)].

C. Chemistry

We will make use of standardized exams prepared by the American Chemical Society, which makes available a wide variety of carefully constructed, nationally normed tests [ACS, 2012]. We will select from among those that are most appropriate for our course offerings, focusing particularly on General Chemistry Form 2009, General Chemistry (Conceptual) Form 2001, and Organic Chemistry Form 2008. By administering such exams before and after instruction in our three primary chemistry courses, we will create a baseline of student learning data upon which to measure and assess future course reforms. Key references for research-based reforms in chemistry include [Gilbert et al., 2002] and [Pienta, Cooper, and Greenbowe, 2005; 2008]. These references includes such evidence-based practices as the Science Writing Heuristic, Peer-Led Team Learning (PLTL), Structured Study Groups, use of Models and Modeling in Chemistry Education, and use of Multimedia Simulations, as well as numerous research results that are used as building blocks for other instructional techniques. Our survey and analysis of current chemistry offerings will explore the degree to which such evidence-based practices are currently being used.

D. Engineering

CTI has a particularly strong faculty representation in engineering education with Profs. Ann McKenna, Odesma Dalrymple, Shawn Jordan, and Adam Carberry all focusing research efforts in that area. We will explore use of several research-based diagnostics in engineering education, e.g., the Thermal and Transport Science Concept Inventory [Miller et al., 2011] and the Materials Concept Inventory developed by ASU Prof. Stephen Krause and his students [Krause and Kelly, 2011].

E. Cross-cutting Curricula

The PI has developed and taught a new course called “Physical Science by Inquiry” (SCN 294) which utilizes one of the most thoroughly validated evidence-based STEM curricula, *Physics by Inqui*ry [McDermott et al., 1996]. This course is targeted at education majors and fulfills the University laboratory science requirement; it is the *only* science course at ASU using evidence-based instructional methods specifically developed and validated for future elementary-school teachers. We will systematically document instructional activities and student learning gains in this course to improve its effectiveness.

VI. Work Plan

Project work will be carried out by the PI and Co-PIs and a full-time graduate research assistant (GRA) who is a Masters or Ph.D. student in STEM or STEM education. The GRA will handle the logistics of test administration and data collection using various diagnostic instruments; PI and Co-PI Zandieh will review course syllabi and instructional materials, observe classes, and interact with faculty. Other STEM faculty will participate by helping to analyze data and adapt evidence-based instructional materials for use in their own classes. Zandieh will review evidence-based instructional materials for possible incorporation in CTI math courses. Project staff will jointly analyze the data and assemble reports for review and discussion with CTI faculty and administrative personnel. The PI will adapt, develop, and implement evidence-based reforms for use in his own physics courses and for potential use in courses taught by other faculty. Project staff will assemble and catalog a resource library of evidence-based materials from diverse STEM fields, initially by reference to [Meltzer and Thornton, 2012], and [Pienta, Cooper, and Greenbowe, 2005; 2008]. They will add suitable annotations and indexing aids, interact with other STEM faculty to make the materials available to them, and provide to the faculty all desired assistance with analysis, adaptation, and implementation of the methods and materials in CTI STEM classes.