PROJECT SUMMARY

Overview:

We will create a national model for improving STEM higher education through a community-based web resource to help physics faculty transform their teaching by incorporating teaching methods and assessments based on research into classroom learning. Historically, higher education struggles to find ways to evaluate and improve instruction. We propose using a 'bottom-up' approach: instead of assessments imposed from accreditation agencies or higher education administration, in our model, discipline-based teachers/researchers develop, share, and aggregate assessment data.

Research-based assessment instruments have had a major impact on physics education reform. They provide universal and convincing measures of student understanding that instructors can use to assess and improve their teaching. The Force Concept Inventory (FCI) has been given to thousands of students worldwide; the use of similar instruments in nearly every subject area of physics is becoming increasingly widespread.

These instruments can transform teaching practice by informing instructors about their teaching efficacy so that they can improve it. At the same time, their widespread use can transform researchers' understanding of the impact of educational transformation by providing large quantities of data that compare teaching practices across a broad range of institutions and student populations. Our preliminary work suggests that physics faculty are eager to use their cognitive resources as scientists to explore big data and compare their students' assessment results to those of other students like their own.

Preliminary work suggests that while the use of these instruments is widespread, most instructors who use them do not know how to interpret the results or how to use them to improve their teaching. Further, because local results are known only to individual instructors, researchers cannot harness the large scale on which these instruments are already given.

We propose to turn the private practice of administering assessment instruments into a community practice of interpreting assessment results in the context of a large community of educators using similar practices in similar settings, comparing results, and using them to transform teaching practices both for individual faculty members and for departments as a whole. We will expand a prototype database developed as part of a current NSF grant (WIDER 1256352) into a community forum and data explorer that allow instructors and researchers to easily upload, discuss, and compare their data in an intuitive, interactive, and beautiful way. The data explorer will feature an intuitive user interface inviting exploration and discovery, interactive one-click analysis tools, a scalable database, and robust data security. This system will be incorporated into the PER User's Guide, an NSF-funded (NSDL 0840853, TUES 1245490) project that provides online resources for physics faculty about research-based teaching methods and assessments.

Intellectual Merit :

An elegantly designed and easily accessible user interface will enable faculty members to engage with their students' assessment data and the national dataset. Hundreds of faculty will better understand the effect of their teaching practices on students' learning and in turn develop further their use of evidence-based teaching methods to improve assessment results and student learning. The database of student assessments will be expanded to include results from smaller colleges and minority-serving institutions that traditionally have not been part of STEM education research. When fully operational, the database will include results from hundreds of colleges and universities.

Broader Impacts :

This data system will be based on prototypes being developed with current funding and will attract more users who will populate the system with an unprecedented amount of assessment data. This will open the doors for physics education researchers and faculty to answer questions about students' learning that were previously inaccessible. We expect that the availability of these data and easy-to-use tools for analysis will encourage faculty adoption of effective teaching methods, which in turn will lead to enhanced student learning. The research team for this project combines experts from physics education research, faculty development, computer security, and data mining and visualization.

Overview

We will create a national model for improving STEM higher education through a community-based web resource to help physics faculty transform their teaching by incorporating teaching methods and assessments based on research into classroom learning. Historically, higher education has struggled to find ways to evaluate and improve instruction. We propose using a 'bottom-up' approach: instead of assessments imposed from accreditation agencies or higher education administration, we are proposing a model in which discipline-based teachers/researchers develop, share, and aggregate assessment data.

Potential: Research-based assessment instruments can transform teaching and research

Research-based assessment instruments[1,2] have had a major impact on physics education reform by providing a universal and convincing measure of student understanding that instructors can use to assess and improve the effectiveness of their teaching. Studies using these instruments consistently show that research-based teaching methods lead to dramatic improvements in students' conceptual understanding of physics.[3] These instruments are already being used on a very large scale: The Force Concept Inventory (FCI)[1], a test of basic concepts of forces and acceleration, has been given to thousands of students throughout the world; the use of similar instruments in nearly every subject area of physics is becoming increasingly widespread. According to a recent survey[4] of faculty who are about to participate in the Workshop for New Faculty in Physics and Astronomy, nearly half have heard of the FCI, and nearly a quarter have used it in their classrooms. The use of these instruments has the potential to transform teaching practice by informing instructors about their teaching efficacy so that they can improve it. At the same time, the widespread use of these standardized instruments has the potential to transform researchers' understanding of the impact of educational transformation by providing large quantities of data that compare teaching practices across a broad range of institutions and student populations. Our preliminary work suggests that physics faculty are eager to use their cognitive resources as scientists to explore big data and compare their students' assessment results to those of other students like their own.

Problem: We are not achieving this potential for transformation

In spite of this potential, no current mechanism supports faculty or researchers in using assessment results to transform their understanding of teaching practice. Our preliminary work suggests that while the use of these instruments is widespread, most instructors who use them do not know how to interpret the results or how to use them to improve their teaching. A recent survey[5] of participants in the Workshop for New Faculty in Physics and Astronomy supports our preliminary work, finding that 45% of participants agreed or strongly agreed with the statement "I find it difficult to know how to interpret the results of research-based assessment instruments." Further, because local results are known only to individual instructors, researchers cannot harness the large scale on which these instruments are already given.

Proposal: Develop a community-based data explorer and assessment resources

We propose to turn the private practice of administering assessment instruments into a community practice of interpreting assessment results in the context of a large community of educators using similar practices in similar settings, comparing results, and using them to transform teaching practices both for individual faculty members and for departments as a whole.

We will expand a prototype database developed as part of a current NSF grant (WIDER 1256352) into a community forum that allows instructors to share and discuss assessment results and teaching practices, and a data explorer that allows instructors and researchers to easily upload, play with, and compare their data to others in an intuitive, interactive, and beautiful way. The data explorer will feature an intuitive user interface inviting exploration and discovery, interactive one-click analysis tools, a scalable database, and robust data security. This system will be incorporated into the PER User's Guide,[6] an NSF-funded (NSDL 0840853, TUES 1245490) project that provides online resources for physics faculty about teaching methods and assessments based on physics education research (PER).

Broader Impacts: An elegantly designed and easily accessible user interface will enable faculty members to engage with their students' assessment data and the national dataset. Hundreds of faculty will better understand the effect of their teaching practices on students' learning and in turn develop further their use of evidence-based teaching methods to improve assessment results and student learning. The database of student assessments will be expanded to include results from smaller colleges and minority-serving institutions that traditionally have not been part of STEM education research. When fully operational, the database will include results from hundreds of colleges and universities.

Intellectual Merit: This data system will be based on prototypes being developed with current funding and will attract more users who will populate the system with an unprecedented amount of assessment data. This will open the doors for physics education researchers and faculty to answer questions about students' learning that were previously inaccessible. We expect that the availability of these data and easy-to-use tools for analysis will encourage faculty adoption of effective teaching methods, which in turn will lead to enhanced student learning. The research team for this project combines experts from physics education research, faculty development, computer security, and data mining and visualization.

Theory of Change

The need for data-driven change

There is substantial evidence that implementing meaningful department-wide assessment practices can have a strong impact on increasing the use of research-based teaching methods. For example, in a study of the results of 10 years of assessment of introductory physics courses using research-based assessment instruments at the University of Colorado (CU), Pollock and Finkelstein[7] report, "Collecting and analyzing these data is good not only for individual course assessments, but also for studying and supporting systematic transformation. *We can use such data to move beyond assessments of a single instructor and a single course to observe the factors that support the widespread adoption and effective implementation of educational practices.* For instance, at CU, the data serve as a mechanism for change. Collecting and reporting these data has become a part of departmental practice. Faculty are privately informed of their performance each semester, and given anonymized versions of these plots to contextualize their performance. While far from perfect, it helps us move beyond the standard end-of-term student evaluation as the sole metric of quality. *We are beginning to couple teaching with learning*" (emphasis added).

There is further evidence that systematic assessment is not only helpful, but also necessary, for meaningful transformation of teaching practice. In a literature review of 191 articles on promoting instructional change in undergraduate STEM education, Henderson et al. report, "Successful strategies focused on disseminating curriculum and pedagogy typically involve more than one of the following components: coordinated and focused efforts lasting over an extended period of time, use of performance evaluation and feedback, and deliberate focus on changing faculty conceptions."[8]

Further, STEM department chairs are motivated by increasing calls to implement effective assessment programs to fulfill accreditation requirements: "Acknowledging the growing consensus that student learning outcomes are the ultimate test of the quality of academic programs, accreditors have also refocused their criteria, reducing the emphasis on quantitative measures of inputs and resources and requiring judgments of educational effectiveness from measurable outcomes."[9] Research has found "evidence of a connection between changes in accreditation and the subsequent improvement of programs, curricula, teaching, and learning in undergraduate programs."[9]

Models for organizational change

Two theories of organizational change guide our work. PER has traditionally relied on a "diffusion" change model wherein researchers develop curricula and disseminate them to a (hopefully) willing crowd of adopters. The researchers usually measure their market penetration in terms of the number of instructors faithfully adopting their methods over time. Preliminary research into faculty change suggests that this model for change is excellent at raising awareness, but poor at increasing the number of long-term users.[8] We use the diffusion model to increase awareness of our project and database through presentations at disciplinary society conferences, the hugely influential Workshop for New Faculty in Physics and Astronomy, and the PER User's Guide.

Another model for organizational change is to empower individual faculty and departments to enact their own changes by enlisting their existing pedagogical expertise, scientific training, and desire to serve their students well. Under this "scholarly teaching model,"[10] instructors are internally motivated to improve their teaching and try new methods, but need help overcoming institutional barriers and adapting research-based teaching and assessment methods to their local contexts. Using data-driven research, an ordinary instructor can take a specific research question, gather evidence, and then disseminate the results. The instructor can incorporate this new knowledge into both improved teaching in the classroom as well as further research questions.[11] Unlike traditional diffusion methods, scholarly teaching does not have a prescribed outcome: by empowering faculty to make the changes which are best for their local context, faculty may choose differently than curriculum designers. However, internally generated changes are more likely to be sustained than externally driven ones.

In interviews with faculty, we have found that many faculty are hungry for data-driven change in their classes, but that the barriers to interpreting the data they already collect are too high. Many instructors are not trained to design and implement changes, and lack the expertise to disseminate the results of changes in their teaching to benefit other, similar, instructors.

We propose using the scholarly teaching model to drive faculty change at diverse institutions. We will harness the transformative power of data-driven assessment to excite faculty to make their own changes, and we support them in interpreting their data and comparing it to their peers through a database-backed, online data explorer which simplifies statistical analysis and guides faculty in interpreting their results. Our database will include dozens of assessments at all levels of the curriculum, allowing faculty assess their students' learning in many different courses and to track the longitudinal effects of reforms. In addition to assessments on content, the database and explorer will contain assessments of attitudes, beliefs, and epistemology, such as the Colorado Learning Attitudes about Science Survey (CLASS)[12] and the Maryland Physics Expectations Survey (MPEX).[13] Through exploration of data from these assessments, the data explorer could encourage faculty to implement teaching methods to address non-content goals.

We expect this project to have a long-term impact on faculty change that will just be beginning by the end of the proposed project. This project will develop and refine an online, data-driven tool for faculty to enact change in their own practice and departments, based on extensive formative assessment including faculty interviews and surveys, user testing, and analysis of usage statistics. Once the tool is fully functional, during the final year of the project, we will measure impact by looking for upward trends in usage, preliminary reports of impact in user testing interviews, and a survey of users. In a future proposal, we will measure the long-term impact on faculty change.

Results from Prior Support

WIDER: EAGER: Increasing faculty use of formative and summative assessment

We began work on a current NSF grant (WIDER 1256352, PIs: McKagan and Sayre, Co-PI: Hilborn, postdoc: Madsen) in February 2013 to develop professional development workshops and online resources around assessment for physics department chairs. One small component of this grant is to develop a prototype database of results from research-based assessment instruments. An NSF grant (TUES 1245490, PI McKagan) funded in April 2013 included interviewing potential stakeholders as part of a needs analysis for this database. We have now conducted this needs analysis, which has informed our plan for developing a database and data explorer that will transform the practices of physics faculty members, physics department chairs and course coordinators, and physics education researchers. Our needs analysis included:

- Interviews with 13 physics faculty members at diverse institutions (potential users of the database)
- Interviews with our Advisory Board, including a chair of a small physics department, a former chair of a large physics department, and a leader in departmental course transformation
- Interviews with developers of other databases of assessment instruments and results
- Consultation with our user interface design specialist (Martinuk) regarding the design of the database
- Consultation with the Kansas State University (KSU) Institutional Review Board (IRB) regarding the protection of human subjects
- Consultation with security expert (co-PI Vasserman) regarding how to ensure security of data and protect anonymity for students, instructors, and institutions
- Consultation with data visualization and informatics expert (co-PI Hsu) regarding how to mine data to help instructors compare to students like theirs and evaluate their teaching
- Consultation with our external evaluator (Henderson), an expert in research on faculty change, regarding how to collect information about faculty teaching practices

The results of these interviews and consultations are discussed below in the section on Products. Based on these results, we have begun work on designing and developing the database. So far, we have:

- Begun designing the user interface for uploading data and for the interactive data explorer.

- Developed a design for a study to test different ways to ask faculty about their teaching practices to determine the best way to collect this information for the database (see supplementary documents).
- Determined the structure of information to be stored in the database, how it will be organized, and what questions we will ask users in order to collect this information (see supplementary documents).
- Developed a prototype of the backend of the database.

We also completed several projects for our current WIDER grant that are not directly related to this one:

- Conducted literature reviews on assessment instruments including FCI, CLASS, and MPEX.
- Designed and implemented a backend for creating assessment instrument guides (see supplementary documents).
- Created a draft assessment instrument guide for the FCI (see supplementary documents).
- Wrote and submitted an article on gender differences on concept inventories to *Physics Review – Special Topics: Physics Education Research* (see supplementary documents).
- Performed a meta-analysis of CLASS and MPEX results that will go into paper on *Top 10 Results of PER* to be submitted to *Physics Today* (see supplementary documents).

By the time this proposal is funded or soon after, we expect to have completed the following additional work on our current WIDER grant, which will provide the groundwork for the current proposal:

- Design and develop front end for assessment instrument guides.
- Write overviews of ~50 assessment instruments and detailed guides of 5-10 assessment instruments.
- Write article on best practices for administering and interpreting the results of research-based assessment instruments, to submit to *The Physics Teacher* and publish on the PER User's Guide.
- Design and develop a prototype system for uploading data into database, basic analysis, and generating reports for faculty comparing their results to results for other students like theirs.
- Write draft overview of IRB and security issues.
- Conduct interviews with physics department chairs.
- Conduct a study to test different ways to ask faculty about their teaching practices and revise database questionnaire based on the results of this study.
- Conduct synthesis research analyzing initial data collected in database.

PER User's Guide

The PER User's Guide[6] is a growing web resource to help physics educators learn about and apply the results of PER in their classrooms, developed by the American Association of Physics Teachers (AAPT) in conjunction with the ComPADRE digital library.[14] Research on faculty change has found that even educators who know about PER and are highly motivated to improve their teaching have trouble finding the information they need to implement PER-based teaching methods effectively.[15] The PER User's Guide is a project designed to improve physics education by translating, summarizing, and organizing the results of PER in an accessible and useful way for busy educators. A previous grant (NSDL 0822342, PI: McKagan) funded the creation of a pilot site and initial user testing, and a current grant is funding ongoing development (TUES 1245490, PI: McKagan)

The pilot site currently contains: overviews of over 50 research-based teaching methods; a list of resources in PER; answers to frequently asked questions and concerns about PER; a guide to what makes PER-based teaching methods work; and a wizard to help instructors find methods to use given their environments and priorities. The wizard allows instructors to filter teaching methods based on criteria such as the instructional environment, the instructor's goals, and various types of research validation.

We have placed our initial focus on PER-based teaching methods because the field of PER has focused heavily on the development and dissemination of specific methods, including curricula, techniques, and resources, as a key mechanism of educational reform.[8] There are also many other important mechanisms and other aspects of PER that are useful to instructors, for example, content-specific teaching suggestions, which are the focus of other pending proposals. The current proposal will allow expansion of the site to include PER-based assessment resources and community tools for faculty.

Cybersecurity

Vasserman is co-PI for the NSF Scholarship for Service Capacity-Building grant (DUE 1129534) for development of "An Innovative Cybersecurity Curriculum for Civilian and Military Workforce" ($299,652, 9/15/2011 -8/31/2013), which has already led to the development of a computer security course for non-traditional students and those with non-STEM backgrounds. The class was offered at Kansas State University in the spring of 2013, and results are being evaluated to further inform development of security education modules which can be incorporated into an existing curriculum. Likewise, the EAGER grant "Education-optional Security Usability on the Internet" ($148,825, 9/1/2012 - 2/28/2013) is funding development of feedback systems to improve the safety of Internet use by people untrained in (or even unaware of) cybersecurity issues. Preliminary results have shown that safe performance is not improved either by drawing further attention to the information presented as part of a security warning, or replacing warning text with graphical information meant to increase risk sensitivity. Research is ongoing.

Vasserman is also a co-PI on "CPS: Synergy: Collaborative Research: Trustworthy Composition of Dynamic App-Centric Architectures for Medical Application Platforms" (CNS 1239543, $880,000, 9/1/2012 - 8/31/2015) and PI on "TWC TTP: Small: Security, Privacy, and Trust for Systems of Coordinating Medical Devices" (CNS 1224007, $482,125, 9/1/2012 - 8/31/2015). These grants support the continued development of the Medical Device Coordination Framework - a high-assurance middleware designed to enable and enforce safe and secure medical device coordination (documentation and code available online: http://mdcf.santos.cis.ksu.edu/download). These grants have funded new work on a domain-specific language for device coordination and real-time reasoning about system capacity and security, individual device functionality, requirements, and performance characteristics (documentation and code available online: https://github.com/mdcf/). Additional cybersecurity research in the medical area is supported by the recent CAREER award (CNS 1253930) for "Safety and Security for Next-generation World-scale Real-time Medical Systems" ($499,020, 3/1/2013 - 2/28/2018). Due to the recent starting date, no results are yet available.

Machine Learning, Probabilistic Reasoning, Data Mining and Visualization

Co-PI Hsu's recent research on heterogeneous information networks, social media, and social networks[16,17] was first facilitated through NSF-supported work on graphical models for gene expression network modeling ("Algorithms for Discovery of Bayesian Network Models of Gene Regulation in Saccharomyces cerevisiae from Microarray Data" (Cooperative agreement 9874732, NSF EPSCOR First Award, 6/10/2002 - 8/9/2003, $75,000) that led to continuing collaborative work with bioinformatics specialists at Kansas State University[18] and to an Office of Naval Research and Department of Homeland Security-sponsored Multimodal Information Access and Synthesis (MIAS) program (2007-2010) with the University of Illinois. The EPSCoR project supported the dissertation research of one doctoral student (Guo)[19,20], and the above projects have supported and engaged two other doctoral students and 10 master's students, including three women among Hsu's advisees[21,22,232425,26].

Products

We are currently funded to build a prototype database with a basic interface for uploading data and to create static reports showing a basic analysis of uploaded data. We expect a beta version of this prototype to be functional and seeded with data by the end of 2013, with user testing and refinement continuing into 2014. The proposed project will build on and expand this work by creating an interactive data explorer, a robust and scalable database, refined and customized assessment reports, and tools for researchers and faculty to mine the database. Our needs analysis has made it clear that while the prototype will be useful, the expanded system outlined in this project will be much more effective at meeting users' needs and inspiring large-scale change through systematic assessments for a wide variety of faculty. Funding this project immediately, rather than after the prototype is completed, will leverage existing funds more efficiently by using them towards the development of the expanded system.

Interactive Data Explorer

We will expand our prototype database to include a large-scale interactive data explorer that allows faculty to easily upload their assessment data and perform intuitive one-click analyses to compare their students' scores to national scores, scores from similar institutions, and their own students' scores from previous years. This one-click analysis will include both the statistical rigor that physicists expect of their data and guides to interpreting the kinds of statistics that education researchers use which are unfamiliar to physicists. An example static visualization is included in Figure 1.

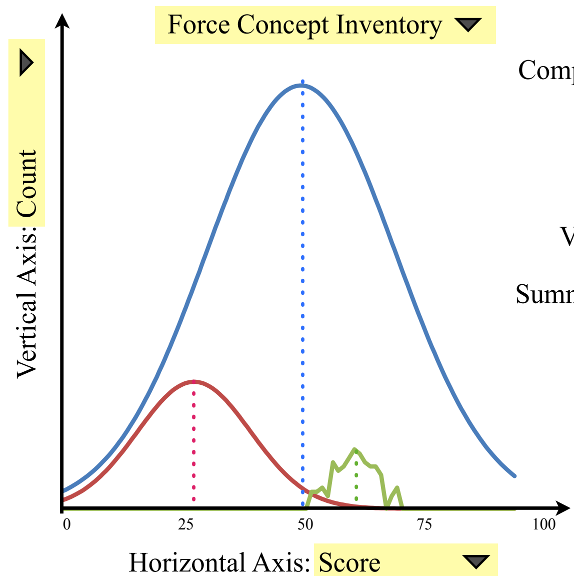

Figure 1: Sample visualization from the interactive data explorer. Users can adjust the comparisons displayed, the vertical and horizontal axes, and the kinds of statistics reported. Where appropriate, we will include error bars and links to tutorials interpreting the results of statistical tests.

The data explorer will include customized tools for three different types of users: individual faculty, department chairs or course coordinators, and researchers. Individual faculty will be able to use it to explore data for their own students, comparing it to averages for their institution, similar institutions, similar classes, and nationwide. Department chairs will be able to use it to compare and analyze results for their whole department, evaluating departmental transformations and comparing the effectiveness of different classes within their departments. Researchers will be able to use it as a tool to explore the effects of a wide variety of factors on aggregate data from across the nation. A preliminary analysis by Kanim[27] has shown that PER does not study a representative sample of the students taking introductory p hysics in this country (Figure 2). Researchers tend to study students at their own institutions, which tend to be large research universities. PER tends to neglect students at 2-year-colleges, liberal arts colleges, and minority-serving institutions, and may be missing issues that are important to learning for some of the students who need it most. Even if faculty at more diverse institutions had the time to comb the research record, they still would not find students like theirs. Our database will allow us to get a more representative sample and help faculty find data for students like theirs by making it easy for any instructor to upload and compare their data.

There are several existing interactive data visualizations which will serve as models for this project. Gapminder[28] is an interactive tool for visualizing various measures of wealth and well-being and comparing them across countries and time. It has a simple intuitive interface that allows users to change the displayed data with the click of a button. This system maps onto our proposed project by replacing country with course, wealth with assessment score, and geographic region with properties of courses or institutions. The Baby Name Voyager[29] is

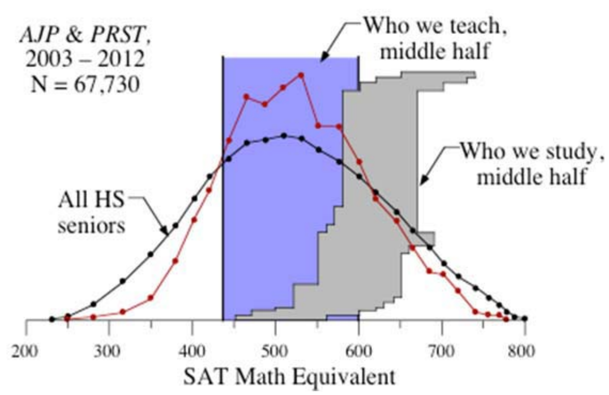

Figure 2: Differences in population between the research record and the national physics-taking student profile

an interactive tool for visualizing the popularity of baby names as a function of time. One of the strengths of the Baby Name Voyager is the way it invites users to discover trends in naming through an intuitive interface and real-time searching. The interface could map to our system by replacing names with student demographics or instructional features and time with score on research-based assessments. Both visualization tools are built on open source software, which we will use for adapting these tools to our own data visualizations.

Results from our interviews support the data explorer:

- Many faculty currently use research-based assessment instruments to assess their own teaching. Some use them to measure the effectiveness of their teaching after making a change in their own teaching methods. Some faculty are interested in assessing their entire physics program and not just single classes. They would like to use assessments to track their students' improvement throughout their college careers. A national database would allow them to compare how well their majors do compared to those at other schools.
- Faculty are interested in comparing their students to their past students, other students at similar institutions, and students at diverse institutions which use similar teaching methods.
- Faculty do not think student evaluations are a good way to measure teaching effectiveness and are interested in having a better way to demonstrate the effectiveness of their teaching.
- Faculty don't know what analyses to do and are excited about the prospect of a system that suggests analyses for them and allows for one-click analysis. They are interested in seeing a variety of complex analyses including error bars and distributions.
- Faculty are excited about the prospect of "playing" with the data and using it to think about their teaching like scientists. For example:

> "My research area is high energy physics and we love lots of data and if we had lots [of student data] we could at least gain insight to certain things, we could make those comparisons."
>
> "If I had unlimited resources this [kind of analysis] is what I would love to do. Of course, then one would be able to experiment: if you do these kinds of activities, do you improve on the following mark or not?...This is basically like a big research project, right: how do you teach this [material]? You would have a really good course after 10 years, but [we're not] set up to do this [testing and revision]. So we try to do it in little bits and pieces."

Interactive Data Explorer: PER data mining and visualization

The data explorer will facilitate applications of data mining[30] to user modeling, adaptation, and personalization tasks:[31,32] student profile clustering,[33,34] classification,[35] and problem selection. Co-PI Hsu is an expert in data mining and visualization, and will lead development and analysis in this respect**.**

These problems are important to solve for users because they will generate interactive and intuitive visualizations of student data for instructors to explore. In specific, we will uses common machine learning techniques[36,37] to provide an interactive visualization[38] of student clusters, descriptive statistics regarding students by category (as assigned by classification or, Bayesian inference, or case-based and analogical reasoning),[39] and associations between student errors. By presenting these statistics visually, instructors will have more direct access to cutting-edge predictive analytics tools based on inductive and analytical machine learning.

Our work with this data is also interesting scientifically. The intellectual merit here is two-fold: first, we will generate new scientific knowledge about how students learn physics as a function of their institutional type, instructional strategies, and demographics. We will also generate new knowledge in the realm of data mining and machine learning. Our work will investigate decision support systems, such as recommender systems for instructors, tutoring, and critiquing. A common recommender system is Netflix (www.netflix.com), which rates new movies for you based on what you enjoy and viewers like you enjoy. A recommender system for this project could recommend instructional strategies for instructors given what the system knows about their students and students like theirs. Other open research problems within this area of intelligent systems include providing functions for causal explanation[40] to instructors or students, automatic relevance determination[41] (also known as feature selection and variable elimination), and its close relatives, feature extraction and construction[42,43] and theory-guided constructive induction.[44]

Robust and scalable database

In order to support one-click analysis that queries data from hundreds or thousands of classes in real time, the prototype database will need to be scaled up and optimized for online analytical processing (OLAP) and basic data warehousing operations. The type of analytical queries that are typical of this project include aggregation queries and multivalued-joins typical of traditional relational databases, but are most often typical of column-oriented and key-value (NoSQL) databases. Scalability requirements include both an increasing number of institutions (from 50 to 150 during the project life cycle, with sustainable long-term growth of 25 per year) with linear growth in the number of courses and students.

Refined and customized assessment reports

In addition to the interactive data explorer, the system will generate static printable pdf reports summarizing instructors' assessment results. These reports can be used by instructors to interpret their own results in order to improve their teaching, to explain what they are doing to colleagues, and to include in teaching portfolios and promotion and tenure reports to demonstrate their teaching efficacy. As part of our current WIDER grant, we will produce two simple versions of these reports (one for instructors and one for course coordinators or department chairs). This proposal will allow us to refine and customize the reports through extensive user testing and incorporate new types of analyses from the new data explorer.

Simple versions of such reports have been produced at the University of Colorado for the FMCE[45] (Pollock) and the E-CLASS[46] (Lewandowski and Zwickl), and faculty have used them in their tenure packages. Both individuals using them and members of the tenure committee have found them to be extremely valuable. In our interviews with faculty and department chairs, both have said that they would use find such reports valuable. Faculty who have implemented reforms, and have found that their colleagues and/or their students are skeptical, were particularly interested in having short reports that could explain the impact of these reforms and how their assessment results compared to national averages. Department chairs on our advisory board said that they would look favorably on receiving such reports in tenure and promotion reports or in job applications.

Community tools

Faculty we interviewed have expressed an interest in connecting with other physics professors who have high assessment scores and/or are using similar teaching methods at similar institutions, in order to discuss their teaching methods and assessment results. A possible model for an online community of faculty connected through an assessment database is PatientsLikeMe (patientslikeme.com), a site in which people can log on, find other people who have the same diseases, find information about the disease, and both contribute and see the results of data on how the disease is progressing and the impact of various treatments. This site provides users with an action research kit for helping patients take control of their own health, and provides researchers with large amounts of data that have led to research published in peer-reviewed medical journals. This site provides a model, with instructors replacing patients and assessment data replacing disease data, for a site in which instructors share uploaded data anonymously, talk to each other through aliases, and learn to improve their assessment and teaching practices through analyzing their own and others' data. In this project we will develop a simplified version of a community based on this model, with a basic forum that allows faculty to create avatars through which they can tell their stories, exchange ideas about teaching and assessment with other faculty, and find faculty in similar situations, while protecting their anonymity with aliases. These community tools have the potential to build synergistically on a TUES Type 2 proposal submitted by PI McKagan in January to build a comprehensive community-driven system for uploading, modifying, rating, and commenting on materials in the database of curricular materials. If that proposal is funded, we will build on it to develop a more extensive community-driven system for this project.

Uploader tool

In interviews we have found that faculty and course coordinators are interested and willing to upload their data to our system, but only if it is extremely easy and if we address their concerns about human subjects protection and security and privacy of their students' data. We will discuss the latter concerns in later sections. We are working closely with our user interface designer (Martinuk) to develop a system that is extremely easy and intuitive to use. We have found in interviews that most faculty imagine a much more complex system on first blush, and when we describe our planned system, they get very excited about it.

Our uploader tool will make it easy and intuitive for instructors and course coordinators to upload data from spreadsheets in nearly any format. Instructors will answer questions about themselves, their classes, and their teaching methods through a web form. We are working carefully to minimize the number of questions instructors need to answer in order to facilitate broad participation. We will use the Carnegie database to classify institutions so that we do not have to rely on self-reports about the properties of their institutions. Instructors will be asked to provide individual student responses to assessment instruments, along with optional information about their students' demographics and backgrounds, by uploading one or more csv or Excel files. An automated system with an intuitive user interface will then guess what the columns represent, and present its guesses for the human uploader to check and correct. This system will draw on best principles of human computer interaction, maximizing efficiency by letting the machine and the humans each do what they do best.

We will develop uploader interfaces for two types of users: individual instructors, who manage and upload assessment data only for their own classes, and department chairs or course coordinators, who manage and upload data for classes taught by many different instructors in their department. Our faculty interviews suggest that currently most departments have assessment data managed by only one or the other of these types of users. However, as the database expands and the use of assessment data for tenure and promotion reports and departmental evaluation increases, it is likely that these two types of users may want to share assessment data with each other, which will require extensive permissions management. We will add permissions and sharing towards the end of the project, after careful study of how different types of users are using the database and their needs for sharing data.

We will provide strong incentives for uploading data through the interactive data explorer and assessment reports described below, which will be available only after faculty have uploaded data. (Users can test drive the data explorer using national data before uploading, but to use it for their own data, they must first upload this data.)

Researcher Tools

The data collected in this database will provide many opportunities for research in physics education in addition to the research on machine learning and data mining detailed in the "Interactive Data Explorer: PER data mining and visualization" section above. The database we develop as part of our current WIDER grant allows us to undertake many exciting research projects, and the expanded database in this proposal will allow us to do many more. In addition, we propose to create tools that allow other researchers to mine the database for their own purposes.

In our discussions with researchers in PER, we have found that many of them are extremely enthusiastic about the possibility of using our database to do research (see letters of support). We will study some of these open questions when the database is sufficiently populated. Examples of unsolved problems in PER that our database will address include:

- By making it easy for any instructor to upload their data, our database will collect a much more representative sample of student data, so that PER studies are no longer restricted to studying students at the institutions of researchers.
- Researchers studying small classes often cannot draw meaningful statistical conclusions from their data due to small numbers. By aggregating data from many similar institutions with small classes, we can collect a sufficiently large sample to conduct high-quality quantitative analysis of small classes.
- Researchers often conduct research on their teaching methods, in whose success they have a vested interest. By anonymizing data and separating the researcher from the instructor, we remove one potential source of bias in the analysis.
- The 1998 Hake study[3] is often cited as evidence that interactive teaching methods lead to dramatic improvements in student learning over traditional lecture, and many studies on individual classes have confirmed the results of this study. There is a need to update this study by including data from the last 15 years and addressing several methodological issues. Our database will provide the data to conduct a comprehensive update to this study and collect more detailed evidence on the impact of teaching practices on student learning.
- Our database will allow us to do comprehensive studies of which characteristics of courses, students, institutions, and instructors matter most for student learning; the size of the data set will also allow us to model the size of each of these effects, to discover interaction effects, and recommend

instructional choices for different courses and populations.
- Our database will allow us to conduct comparisons and correlations between different research-based assessment instruments, for example, correlating scores on conceptual assessments with assessments of beliefs and attitudes.

Addressing Potential Obstacles

Building a robust data set

The success of this project depends on our ability to seed the database with enough initial data to which potential users can compare their own students' scores. In order to be useful for faculty at all kinds of institutions teaching all kinds of classes, the data, taken as a whole, must be as representative as possible for the population of people who study physics. We are undertaking several measures in order to ensure that we have a large and robust data set.

First, inasmuch as data can be gleaned from published works, we will enter published results into the database. Where published studies do not contain sufficient detail, we will contact authors of published studies and ask for their raw data to include in the database. Thus the database will be seeded with the already extensive results of published studies. We will build tools that allow separating out published data to make easy comparisons between published and new data.

Second, we have asked faculty from institutions that have collected large quantities of assessment data and/or serve underrepresented populations to commit to uploading their data to a beta version of the database in order to seed it. In the supplementary documents we have included letters of support from 12 faculty who have committed to upload data for their departments, including 7 from institutions that have been collecting assessment data for more than 10 years, 1 from a 2-year college, 3 from minority-serving institutions, and 3 from small liberal arts colleges.

Collecting accurate information about faculty teaching practices

For our database to be useful for comparing the effectiveness of different teaching methods, we must collect accurate information from faculty about their teaching practices. This is a notoriously difficult problem; research has shown that when faculty report that they use a particular PER-based teaching method, their practices often bear little resemblance to the practices suggested by the developers of that method.[47] The ideal way to collect this information would be for trained researchers to observe classrooms. Since this is not feasible for a large-scale collection of archival data, we are instead working to determine the most accurate way to collect this data from faculty self reports. We have reviewed several previous surveys of faculty teaching practices, including the 16-page survey[48] that formed the basis of the Hake study,[3] two surveys of physics faculty,[5,49] a survey of engineering faculty,[50] and a survey of biology faculty.[51] Based on these surveys and consultation with our external evaluator (Henderson), we have determined that asking specific questions about what happens in the classroom with objective numerical answers (e.g. "What fraction of class time do your students spend talking to each other?") is likely to result in more accurate responses than questions about which teaching methods faculty use (e.g. "Do you use interactive engagement?") We have designed a study (see supplementary documents) to use faculty surveys to compare different ways of asking about teaching practices. We expect to have completed this study before the proposed project begins.

Human subjects protection

Many faculty, particularly physics education researchers, are naturally concerned that our database must comply with human subjects regulations, and are reluctant to upload data without assurance that they have permission to do so. We have been in contact with the leaders of other projects that include large-scale multi-institution collection of aggregate student data, including the Wabash National Study[52] and the Geoscience Concept Inventory database,[53] and both provide clear precedents for getting Institutional Review Board (IRB) permission to collect such data. We have also been in discussion with the KSU IRB, and based on these discussions, we expect to receive permission to collect archival data from faculty from any institution and to keep this data in perpetuity. Because all student names and numbers will be anonymized and the data will be reported only in aggregate, students are not considered research subjects in this study. Faculty are considered research subjects, and will need to submit an online consent form, which will be seamlessly integrated into the upload process.

We have found that the biggest obstacle in terms of human subjects protection is not getting permission for this project, but the *perception* by much of the PER community that human subjects regulations will make getting permission hard or impossible. We intend to address such perceptions by helping other people with their IRBs, and doing user education to support people who are not IRB-aware in protecting the anonymity and confidentiality of their students. We will set up our system to automatically protect students, faculty, and institutions, even when those uploading the data do not take precautions to do so.

Because IRBs at different institutions interpret regulations very differently, we recognize that not all institutions will be satisfied with the KSU IRB's decision that it has the authority to regulate data from other institutions. To address concerns of faculty at such institutions, we will develop sample IRB forms that they can use to get permission from their own institutions' IRBs to upload data, as well as documents and answers to frequently asked questions addressing their concerns. Based on our faculty interviews and discussions with potential uploaders, we suspect that the fraction of institutions where the KSU IRB approval will be insufficient is substantially smaller than is commonly believed.

Security and privacy

We have planned many measures to protect the security of all data and the privacy of students, instructors, and institutions. These efforts will be led by co-PI Vasserman, who has extensive experience with database security. Only individual instructors and course coordinators will be able to see individual data for their students, themselves, or their institutions. Privacy settings will allow instructors to choose whether course coordinators or others in their own department can see data they have uploaded, and vice versa. Other users will see average aggregate data for types of institutions, instructors, or students, but not individual data for any particular institution, instructor, or student. We will have a system in place to protect anonymity by preventing queries that would reveal individual identities because the number of similar entries in the database is too small. We will make simple descriptions of our security measures, including frequently asked questions, that are easily accessible to faculty considering uploading data in order to address their concerns.

When dealing with student data, privacy is a natural concern, especially with ongoing and interactive access by a large number of researchers. While the security of databases (preventing inference of data which was not directly queried) is unsolvable in the general case,[54] we can take numerous precautions to minimize if not eliminate the danger of data disclosure. These steps not only enhance security and privacy but also increase usability and uptake, since instructors and researchers will be more likely to trust databases with built-in protection.

First and foremost, data must be secure locally on the database servers. This means servers with the latest anti-virus and anti-malware software, intrusion detection systems, and limited physical access to trusted project personnel. Sensitive data should be encrypted on disk (using whole-disk encryption), and it may be feasible to deploy in-memory encryption for maximum security assurance. All relevant activity (e.g. logins, login failures, database operations performed by local users, etc.) must be logged, along with the responsible party to preserve accountability and prevent incidents.[55,56]

Data within the database will be anonymized or at least de-identified (replacing personally identifiable fields such as names and student IDs with consistent but random values). For instance, a student name is transformed into an alphanumeric identifier which cannot be tied to the original user due to the nature of that transformation (one-way hash functions). While the new value appears unrelated to the original name, the student can still be tracked as an individual – every time data for the same student is added, the same transformation is applied, and the same alphanumeric value is derived. The data stored in the database after processing allows analysis of within-student, between-student, and aggregate statistics without ever exposing personal identifying information. Likewise, demographic data will be sanitized to reduce the risk of personal identification of students[54,57] – this can be achieved by adding a small level of noise or simply not exposing that particular field except in aggregate queries to the database. For example, a query which would return a unique result, e.g. one individual, can simply be refused.

Since the database will be centrally administered and login information recorded and controlled, we can at least alleviate the general inference impossibility problem by refusing to respond to database queries which would yield an unacceptably small number of results (potentially allowing for identification). Some may attempt to bypass such protections by using multiple accounts,[54] but we minimize potential damage using additional safety measures, such as limiting the information which can be derived from queries not only by individual accounts but also by all accounts within an institution. Noise is added to responses in all queries, preserving statistical properties and maximizing scientific value while still protecting data from advanced inference, as in medical research and census data reporting.[57,58] The nature and amount of that noise depends on the individual fields involved in the database query. This system will be based on the U.S. Census, which contains large quantities of sensitive personal information and allows researchers to query the contents, but refuses queries for which there is not a large enough anonymity set to protect the anonymity of the data.

Work Plan

Work on this project proceeds through four phases and across five avenues (Figure 3):

In the first phase ("Pre-development"), the bulk of our efforts are in designing the database, uploader tool, and prototype assessment reports. During this phase, we will solicit data from specific uploaders (see Supplementary Documents for letters of commitment) and convert data from the prototype database in order to seed the new database and fuel development in the second phase.

In the second phase ("Closed-alpha"), we will invite instructors and researchers to interact with all of the parts of our system. Through intensive user testing including both interviews and analysis of interactions with the system, we will optimize the database for common queries, improve the uploader tool, develop more intuitive and informative visualizations, and incorporate those visualizations into customized assessment reports.

In the third phase ("Open-beta"), the system will be sufficiently stable, the features sufficiently rich, and the data in the database sufficiently robust, that we will open it to the public. In this phase, the bulk of our effort shifts from developing radically new tools to refining existing features and conducting research on the contents of the database.

In the fourth phase ("Release"), our development efforts will be minor, but our research and dissemination efforts will continue in full force.

Project Staff

Sarah "Sam" McKagan (PI) leads the design and development of the uploader tool and customized assessment reports. She will also ensure that the products of this project are well-integrated into the PER

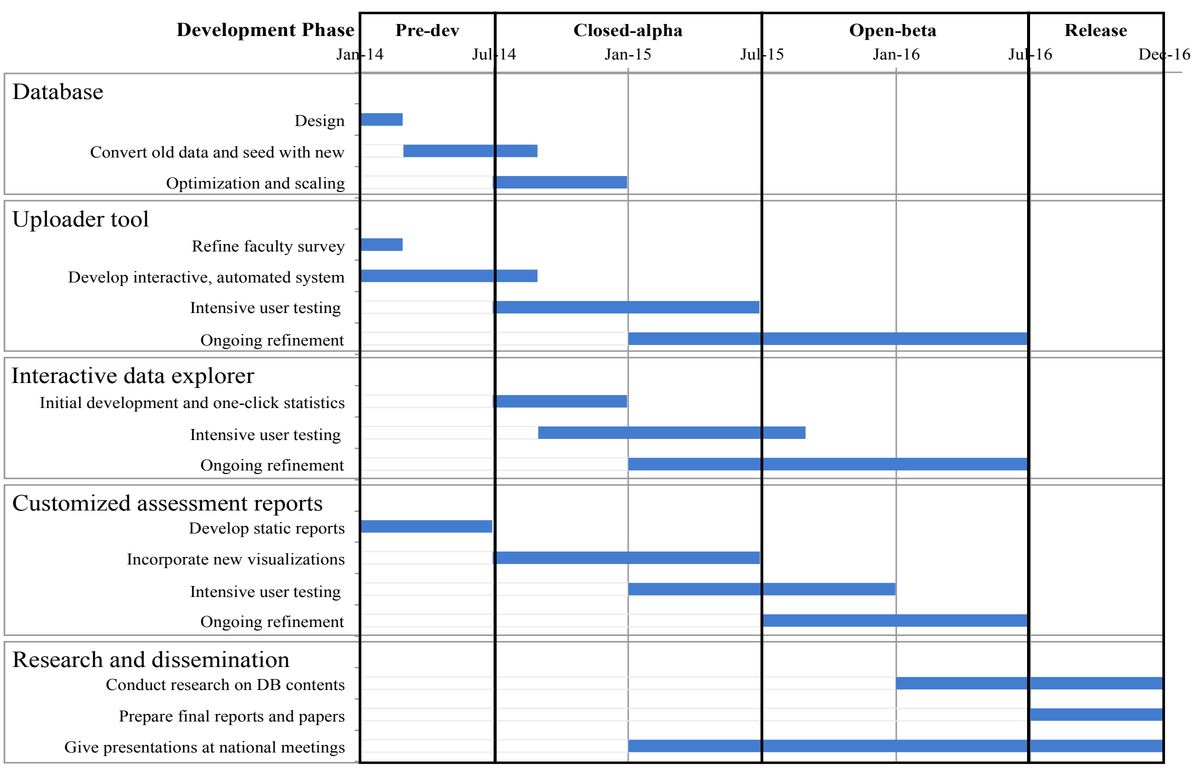

Figure 3: Project timeline

User's Guide. She is the creator and editor of the PER User's Guide for the American Association of Physics Teachers (AAPT), and an education research consultant for Seattle Pacific University, Augsburg College, and the University of Colorado. As a developer of online resources to help faculty learn about PER-based teaching methods, she has observed classrooms, interviewed developers and adopters, and developed expertise in a wide variety of research-based teaching and assessment methods.

Eleanor Sayre (PI) will lead efforts to populate the database with data, conduct research on the contents of the database, and ensure that the data contained therein is both representative of the physics-taking population and the diversity of schools which offer physics courses. Additionally, she leads the efforts to develop and implement the one-click statistics portion of the interactive data explorer. She is an Assistant Professor of Physics at Kansas State University (KSU). Before arriving at KSU, Dr. Sayre led departmental assessment efforts at Wabash College, focusing on how PER assessment methods can be adapted to the small liberal arts setting, and how results can inform departmental reviews. Dr. Sayre developed the Response Curves Methodology which measures student learning and forgetting on the scale of days or weeks in large-enrollment classes.

Adrian Madsen (co-PI) will conduct extensive interviews with existing and prospective users (faculty, course coordinators, department chairs, and PER researchers) to guide development of features and ensure that the project as a whole is meeting the needs of the community. She will also coordinate with PI McKagan to develop customized assessment reports, and with PI Sayre to conduct research on the contents of the database. Madsen is an Assistant Editor at the PER User's Guide. Her past research includes an analysis of gender results on concept inventories, and eye-tracking studies to determine how visual salience affects understanding of physics diagrams.

Robert Hilborn (co-PI) will coordinate the work of this project with professional society efforts in faculty professional development and instructional reform, and will assist with advertising the data explorer to the AAPT community. Hilborn is Associate Executive Officer of AAPT. He is the leader of the Physics and Astronomy New Faculty Workshops that have introduced over 1,600 new physics and astronomy faculty members to the latest science pedagogy and the research that supports that pedagogy. He has also served as staff organizer for the Physics Department Chairs Conference hosted jointly by the APS and AAPT. He led the National Task Force on Undergraduate Physics and its SPIN-UP study of thriving undergraduate physics programs.

Eugene Vasserman (co-PI) will work with us to ensure that our system will protect the security of all data and the privacy of students, instructors, and institutions. He will anticipate and address ongoing and emergent security and privacy concerns. He is an assistant professor of Computing and Information Sciences at Kansas State University. His research is in distributed system security, privacy and anonymity, low-power and pervasive computing, and applied cryptography.

William Hsu (co-PI) will lead the development team to generate the database, uploader tool, and interactive data explorer. He will coordinate with McKagan, Madsen, and Martinuk about users' needs, with Sayre about one-click statistics, and conduct research on applying machine learning to data cleaning, integration, mining, and visualization tasks, using the database as an instance of big data. Hsu is an associate professor of Computing and Information Sciences at Kansas State University. His research is in information visualization techniques for big data, automated recommender systems, probabilistic reasoning, and machine learning.

Mathew "Sandy" Martinuk will serve as a user interface design consultant, working with the PIs on the design of the user interface for the backend and frontend of the data explorer and uploader tool to ensure that they incorporate best-practice design principles, and overseeing user testing of these resources. Martinuk is a User Experience Designer with Theresa Neil Interface Design, a prominent design firm whose recent clients include PayPal, Forbes, and Pearson. Martinuk (along with McKagan) is a key member of the team designing the new *Teach with PhET* website for the PhET Interactive Simulations project, which will be launched in late summer 2013. He has also served as a user interface designer on two previous grants for the PER User's Guide, and is in the process of designing assessment guides and a basic assessment database. His specialty is leveraging his background as a physics education researcher to create educational software user experiences that speak naturally to the needs of educators and students.

Evaluation Plan

We will evaluate the success of our project based on the number of faculty using the system, the robustness and representativeness of the data included in the database, the quality of users' experience, the degree to which use the data explorer affects faculty assessment and teaching practices, and the degree to which our project is leading to research products.

Use of system and robustness and representativeness of the data included in the database

Throughout the project, we will collect and analyze usage statistics to determine how many users are accessing the database and what they are doing with it. Important measures of success will be the quantity of data uploaded to the database, the robustness of that data, and the degree to which it is representative of the population of students taking physics. Figure 4 below outlines the quantity of data that we expect to have in the database at each phase of development. The values for the closed alpha stage are estimates based on the descriptions of the quantity and type of assessment data faculty outlined in their letters of support (see supplementary documents), assuming about 90 students/class. Institutions with large classes are oversampled in this stage. In the open beta phase, we will solicit assessment data from smaller institutions including liberal arts schools and two-year colleges, and we estimate adding data for 600 classes with an average size of 50 students/class across 60 institutions. At release, we will ensure that the database reflects the population of students taking physics. About 50% of physics students in the US are enrolled in universities granting advanced degrees and 50% of students are enrolled at two-year or four-year colleges. [59],[60] . Based on this distribution of institution type and their associated average class sizes, we estimate adding data for 1000 classes with an average size of 70 students per class at the release stage. At steady state, we expect about 100 instructors to upload data each year from about 200 classes with an average size of 70 students per class. Published results will also be included in the database for comparison purposes. These usually contain summary data only.

We will conduct user testing during closed-alpha and open-beta stages to ensure that faculty are correctly interpreting the questions asked by the uploader tool and that their responses are meaningful. We will refine the interface based on this user testing to make it as easy as possible for faculty to enter robust and meaningful data. At each phase, we will compare the database to the population of students taking physics (see Figure 1), and launch campaigns to target specific underrepresented populations.

	Published Results	Closed Alpha	Open Beta	Release	Steady State
Institutions	20	12	75	150	25/yr
Course coordinators	17	20	70	140	10/yr
Instructors	100	115	350	700	100/yr
Classes	400	450	1150	2200	210/yr
Students	27000	40000	75000	150000	15000/yr

Figure 4: Anticipated quantity of data at each stage of development. Data for closed alpha, open beta and release phases are cumulative. Summary data from published results will be included for comparison purposes.

Quality of users' experience

We will work carefully with our user interface consultant (Martinuk) to design a system that is intuitive, useful, and easy. We will conduct formative assessment of the system throughout the project in the form of rigorous user interface testing with faculty at diverse institutions, using an iterative process of testing and redesign. In addition, our user interface evaluator (Redish) will periodically review the interface design to ensure that it follows best practices for usability.

Degree to which use of the data affects faculty assessment and teaching practices

Throughout the project, we will learn about how the use of the data explorer impacts faculty assessment and teaching practices through user testing. During the final year of the project, we will conduct a survey of faculty with accounts on our system to find out how it has impacted their teaching and assessment practices. Because the data explorer will be fully functional only at the end of our project, we expect the time scale of the impacts of the data explorer to be longer term than the time scale of the proposed project. In a future proposal, we will study the long-term effects through interviews, classroom observations, and ethnographic studies of faculty who use the data explorer. In this proposal, we expect usage statistics, user testing, and our faculty survey to show an upward trend in database usage and indications that this usage is beginning to have an impact on faculty practice.

Degree to which project is leading to research products

This work leads to four types of research products in the form of articles in peer-reviewed journals: (1) We will use the data in the database to conduct our own research studies on how people learn physics and the demographic, institutional, and instructional factors that affect learning. (2) We will use the results of our usage statistics, user testing, and faculty survey to conduct research on faculty change, exploring how use of the data explorer impact faculty teaching and assessment practices. (3) Co-PI Hsu will use the database to conduct research on visualization and learning analytics. (4) We expect other researchers to take advantage of our researcher tools to access and use the contents of the database for their own research projects, such as the proposed project by Dr. Kapon (see letter of support) to analyze patterns of mistakes in FCI pretests. To ensure that our work is effective, we will conduct frequent checks with members of our evaluation board, which includes experts in several key areas of the project:

Charles Henderson is a physics professor at Western Michigan University whose research focuses on understanding and improving the historically slow incorporation of research-based instructional reforms into college-level STEM courses. Research projects have attacked this problem from several directions: in-depth studies of faculty attempting to change their instruction[61], surveys and interviews with faculty related to their interaction with research-based instructional strategies[15,49,62], and an interdisciplinary literature review of change strategies in higher education[8]. His work was cited prominently in a major initiative recently announced by the Association of American Universities (AAU) to improve the quality of undergraduate teaching and learning at its member institutions[63], and has resulted in invited presentations at national panels and commissioned white papers[64].Henderson will serve as an overall external evaluator, periodically reviewing the progress of work and offering suggestions for improvements to both research and resources based on the results of his research and his evaluation of the strength of the project overall. He will also advise us on our interactions with faculty. He will meet regularly with the PIs remotely and/or in person at national meetings; review the overall progress of the grant including the timeline and goals; examine and advise on the research study designs, methodologies, and findings; review and advise on the resource design, development, content, and dissemination.

Colleen Megowan is the president of the American Modeling Teachers Association (AMTA), an active community of research-informed high school teachers engaged in the practice of Modeling Instruction, which owns and manages the FCI and several other widely used concept inventories. She is leading the development of an assessment database for high school teachers, and will work closely with our team to integrate the two databases. Using her expertise with building the modeling community, she will advise us on the growth process of building such a community and evaluate the community aspects of our projects.

Janice (Ginny) Redish is a world-renowned specialist in user experience research and design, and the author of *Letting Go of the Words—Writing Web Content that Works*. She will evaluate the design and usability of the web interface for the data explorer, uploader tool, community tools, and researcher tools. She will assist with issues of content and usability including developing personas, doing task analysis, organizing and writing instructions and messages, testing the site against best practices for usability.

Mikhail Voloshin is a software engineer at Google with substantial experience building back-end data storage, front-end client access, and administrative maintenance for many projects. He will advise us and evaluate our work with bulk data management, database design, and software architecture.

Dissemination and Publicity

This project will include a large-scale publicity campaign, with guidance from the AAPT marketing director, to develop awareness and interest among physics educators. The newsletters and electronic announcements of the AAPT and APS will be used to disseminate the work to their members, which represent a large fraction of college physics faculty. The PER Leadership and Organizing Council (PERLOC), the representative body of the PER research community, will advertise the PER User's Guide, including the data explorer, to its constituents, encouraging researchers to contribute and use it as a resource for educators with whom they work. The PIs will give presentations about the data explorer at national and local meetings of the AAPT and APS, and the data explorer will be featured in workshops on the PER User's Guide and ComPADRE. The AAPT will feature the data explorer in the many workshops it oversees, including the Physics and Astronomy New Faculty Workshops (which reach about 50% of new physics and astronomy tenure-track hires in the U.S.), the new Experienced Faculty Workshops, physics department chairs meetings, and SPIN-UP workshops.

References

[1] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach.* **30**, 141 (1992).
[2] R. Beichner, "Testing Student Interpretation of Kinematics Graphs," *Am. J. Phys.* **62**, 750 (1994); R. K. Thornton and D. R. Sokoloff, "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); Paula Vetter Engelhardt and Robert J. Beichner, "Students' Understanding of Direct Current Resistive Electrical Circuits," *Am. J. Phys.* **72**, 98 (2004); D. P. Maloney, T. L. O'Kuma, C. J. Hieggelke, and A. Van Heuvelen, "Surveying Students' Conceptual Knowledge of Electricity and Magnetism," *Am. J. Phys.* **69**, S12 (2001); C. Singh and D. Rosengrant, "Multiple-choice Test of Energy and Momentum Concepts," *Am. J. Phys.* **71**, 607 (2003); L. Ding, R. Chabay, B. Sherwood, and R. Beichner, "Evaluating an Electricity and Magnetism Assessment Tool: Brief Electricity and Magnetism Assessment," *Phys. Rev: ST PER* **2**, 010105 (2006).
[3] R.R. Hake, "Interactive-engagement vs. Traditional Methods: A Six-thousand-student Survey of Mechanics Test Data for Introductory Physics Courses," *Am. J. Phys.* **66**, 64-74 (1998).
[4] C. Henderson, private communication.
[5] M. H. Dancy, C. Henderson, and J. Smith, "Understanding Educational Transformation: Findings From a Survey of Past Participants of the Physics and Astronomy New Faculty Workshop," submitted to *PERC Proceedings 2013*.
[6] http://perusersguide.org
[7] S. J. Pollock and N. D. Finkelstein, "Impacts of Curricular Change: Implications from 8 Years of Data in Introductory Physics," *PERC Proceedings 2012*, AIP Press. Melville NY (2013).
[8] C. Henderson, N. Finkelstein, and A. Beach, "Beyond Dissemination in College Science Teaching: An Introduction to Four Core Change Strategies," *Journal of College Science Teaching* **39** (5), 18-25 (2010).
[9] J. F. Volkwein, L. R. Lattuca, B. J. Harper, and R. J. Domingo, "Measuring the Impact of Professional Accreditation on Student Experiences and Learning Outcomes," *Research in Higher Education* **48** (2), 251-282 (2007).
[10] M. Borrego and C. Henderson, "Theoretical Perspectives on Change in STEM Higher Education and their Implications for Engineering Education Research and Practice," submitted to *Journal of Engineering Education*.
http://homepages.wmich.edu/~chenders/Publications/2014BorregoJEEPerspectivesonChange.pdf
[11] G. Mettetal, "The What, Why and How of Classroom Education Research," *Journal of the Scholarship of Teaching and Learning* **2**(1) (2001).
[12] W. K. Adams, K. K. Perkins, N. Podolefsky, N. D. Finkelstein, and C. E. Wieman, "A New Instrument for Measuring Student Beliefs about Physics and Learning Physics: The Colorado Learning Attitudes about Science Survey," *Phys. Rev: ST PER* **2**, 010101 (2006).
http://link.aps.org/doi/10.1103/PhysRevSTPER.2.010101
[13] E. F. Redish, J. M. Saul, and R. N. Steinberg, "Student Expectations in Introductory Physics," *Am. J. Phys.* **66**, 212 (1998). http://dx.doi.org/10.1119/1.18847
[14] http://compadre.org
[15] C. Henderson and M. Dancy, "Physics Faculty and Educational Researchers: Divergent Expectations as Barriers to the Diffusion of Innovations," *Am. J. Phys.* **76** (1), 79-91 (2008).
[16] W. H. Hsu, J. P. Lancaster, M. S. R. Paradesi, and T. Weninger, "Structural Link Analysis from User Profiles and Friends Networks: A Feature Construction Approach," in N. S. Glance, N. Nicolov, E. Adar, M. Hurst, M. Liberman, & F. Salvetti (Ed.), *Proceedings of the 1st International Conference on Weblogs and Social Media (ICWSM 2007)*, (pp. 75-80). Boulder, CO, USA (2007).
[17] D. Caragea, V. Bahirwani, W. Aljandal, and W. H. Hsu, "Ontology-Based Link Prediction in the LiveJournal Social Network," *Proceedings of the 8th Symposium on Abstraction, Reformulation and Approximation (SARA 2009)* (2009).
[18] M. S. R. Paradesi, D. Caragea, and W. H. Hsu, "Structural Prediction of Protein-Protein Interactions in *Saccharomyces cerevisiae*," *Proceedings of IEEE 7th International Symposium on BioInformatics and BioEngineering (BIBE 2007)*, (pp. 1270-1274). Boston, MA, USA (2007).

[19] H. Guo, *Algorithm Selection for Sorting and Probabilistic Inference: A Machine Learning-based Approach* (Ph.D. dissertation). Kansas State University, Department of Computing and Information Sciences, Manhattan, KS, USA (2003).
[20] H. Guo and W. H. Hsu, "A Machine Learning Approach to Algorithm Selection for NP-hard Optimization Problems: A Case Study on the MPE Problem," *Annals of Operations Research, Special Issue on Stochastic Search Algorithms*, 156(1), 61-82 (2007).
[21] S. Volkova and W. H. Hsu, "Computational Knowledge and Information Management in Veterinary Epidemiology," in C. C. Yang, D. Zeng, K. Wang, A. Sanfilippo, H. H. Tsang, M.-Y. Day, . . . H. Chen (Ed.), *Proceedings of the 8th IEEE International Conference on Intelligence and Security Informatics (ISI 2010)*, May 23-26, 2010, Vancouver, BC, Canada (pp. 120-125). IEEE Press (2010).
[22] S. Volkova, D. Caragea, W. H. Hsu, J. Drouhard, and L. Fowles, "Boosting Biomedical Entity Extraction by using Syntactic Patterns for Semantic Relation Discovery," *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2010)* (pp. 272 - 278). Toronto, ON, Canada: IEEE Press (2010).
[23] S. O Volkova, "*Entity Extraction, Animal Disease-related Event Recognition and Classification from Web*" (M.S. thesis), Kansas State University, Department of Computing and Information Sciences, Manhattan, KS, USA (2010). Retrieved July 2, 2013, from http://hdl.handle.net/2097/4593
[24] S. Roy Chowdhury, "*Mathematical Models for Prediction and Optimal Mitigation of Epidemics"* (M.S. thesis), Kansas State University, Department of Electrical Engineering and Computer Engineering, Manhattan, KS, USA (2010). Retrieved July 2, 2013, from http://hdl.handle.net/2097/3874
[25] S. Volkova, D. Caragea, W. H. Hsu, and S. Bujuru, "Animal Disease Event Recognition and Classification," *Proceedings of the First International Workshop on Web Science and Information Exchange in the Medical Web (MedEx 2010)*, April 30, 2010, Raleigh, NC, USA (2010).
[26] S. Bujuru, "*Event Recognition in Epizootic Domains"* (M.S. thesis), Kansas State University, Department of Computing and Information Sciences, Manhattan, KS, USA (2010). Retrieved July 2, 2013, from http://hdl.handle.net/2097/7070
[27] S. Kanim, private communication
[28] http://www.gapminder.org/world
[29] http://www.babynamewizard.com/voyager
[30] C. Romero, "Educational Data Mining: A Review of the State of the Art". *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Review* **40**(6), 601-618 (2010).
[31] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Effective Personalization based on Association Rule Discovery from Web Usage Data," in R. H.-L. Chiang, & E.-P. Lim (Ed.), *Proceedings of the 3rd International Workshop on Web Information and Data Management (WIDM 2001)* (pp. 9-15). New York, NY, USA: ACM Press.
[32] R. S., Baker, Z. A. Pardos, S. M. Gowda, B. B. Nooraeil, and N. T. Heffernan, "Ensembling Predictions of Student Knowledge within Intelligent Tutoring Systems," in J. A. Konstan, R. Conejo, J. L. Marzo, & N. Oliver (Ed.), *Proceedings of the 19th International Conference on User Modeling, Adaption and Personalization (UMAP 2011): Lecture Notes in Computer Science Volume 6787* (pp. 13-24). New York, NY, USA: Springer (2011).
[33] L. Razzaq, R. W. Maloy, S. Edwards, D. Marshall, I. Arroyo, and B. P. Woolf, "4MALITY: Coaching Students with Different Problem-Solving Strategies Using an Online Tutoring System," in J. A. Konstan, R. Conejo, J. L. Marzo, & N. Oliver (Ed.), *Proceedings of the 19th International Conference on User Modeling, Adaption and Personalization (UMAP 2011): Lecture Notes in Computer Science Volume 6787* (pp. 359-364). New York, NY, USA: Springer (2011).
[34] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques* (3rd ed.), San Francisco, CA, USA: Morgan Kaufmann (2011).
[35] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.), San Francisco, CA, USA: Morgan Kaufmann (2011).
[36] T. M. Mitchell, *Machine Learning*, New York, NY, USA: McGraw-Hill (1997).
[37] P. Domingos, "A Few Useful Things to Know about Machine Learning," *Communications of the ACM* **55**(10), 78-87 (October, 2012).
[38] R. Mazza, and V. Dimitrova, "Generation of Graphical Representations of Student Tracking Data in Course Management Systems," in E. Banissi, M. Sarfraz, J. C. Roberts, & B. Loften (Ed.), *Proceedings of*

the 9th International Conference on Information Visualisation (IV 2005) (pp. 253-258). New York, NY, USA: IEEE Press (2005).
[39] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *Artificial Intelligence Communications*, **7**(1), 39-59 (1994).
[40] T. Verma, and J. Pearl, "An Algorithm for Deciding if a Set of Observed Independencies has a Causal Explanation," in D. Dubois, & M. P. Wellman (Ed.), *Proceedings of the 8th Annual Conference on Uncertainty in Artificial Intelligence (UAI 1992)* (pp. 323-330). San Mateo, CA, USA: Morgan Kaufmann (1992).
[41] R. Kohavi and G. H John, "Wrappers for Feature Subset Selection," *Artificial Intelligence* **97**(1-2), 273-324 (1997).
[42] M. L. Raymer, W. F. Punch, E. D. Goodman, P. C. Sanschagrin, and L. A. Kuhn, "Simultaneous Feature Extraction and Selection Using a Masking Genetic Algorithm," in T. Bäck (Ed.), *Proceedings of the 7th International Conference on Genetic Algorithms (ICGA 1997)* (pp. 561-567). San Francisco, CA, USA: Morgan Kaufmann (1997).
[43] I. Guyon and A. Elisséeff, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research*, **3**(Mar), 1157-1182 (2003). Retrieved July 2, 2013, from http://jmlr.org/papers/special/feature03.html
[44] S. K. Donoho and L. A. Rendell, "Rerepresenting and Restructuring Domain Theories: A Constructive Induction Approach," *Journal of Artificial Intelligence Research* **2**, 411-446 (1995).
[45] R. K. Thornton and D. R. Sokoloff, "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). http://dx.doi.org/10.1119/1.18863
[46] Benjamin M. Zwickl, Noah Finkelstein, and H. J. Lewandowski, "Development and Validation of the Colorado Learning Attitudes about Science Survey for Experimental Physics," *PERC Proceedings 2012*, AIP Press. Melville NY (2013). http://link.aip.org/link/doi/10.1063/1.4789747
[47] C. Turpen, M. Dancy, and C. Henderson, "Faculty Perspectives on Using Peer Instruction: A National Study," *PERC Proceedings 2010*, AIP Press. Melville NY (2010). http://dx.doi.org/10.1063/1.3515235
[48] http://www.physics.indiana.edu/~sdi/SurveyForm03201997.pdf
[49] C. Henderson and M. H. Dancy, "Impact of Physics Education Research on the Teaching of Introductory Quantitative Physics in the United States," *Phys. Rev: ST PER* **5**, 020107 (2009). http://link.aps.org/doi/10.1103/PhysRevSTPER.5.020107; http://prst-per.aps.org/multimedia/PRSTPER/v5/i2/e020107/e020107_app.pdf
[50] M. Borrego, S. Cutler, M. Prince, C. Henderson, and J. Froyd, "Fidelity of Implementation of Research-Based Instructional Strategies (RBIS) in Engineering Science Courses," *Journal of Engineering Education* (in press).
[51] T. M. Andrews, M. J. Leonard, C. A. Colgrove, S. T. Kalinowski, "Active Learning not Associated with Student Learning in a Random Sample of College Biology Courses," *CBE Life Sci. Educ.* **10**, 394–405 (2011). http://www.lifescied.org/content/10/4/394.full; http://www.lifescied.org/content/suppl/2011/11/17/10.4.394.DC1/CombinedSupMats4.pdf
[52] C. F. Blaich and K. S. Wise, "From Gathering to Using Assessment Results: Lessons from the Wabash National Study," NILOA Occasional Paper No. 8, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment (2011). http://www.learningoutcomeassessment.org/documents/Wabash_000.pdf; http://www.liberalarts.wabash.edu/study-research/
[53] J. C. Libarkin and J. P. Kurdziel, "Research Methodologies in Science Education: Human Subjects and Education Research," *Journal of Geoscience Education* **52**, 199-203 (2004); http://geoscienceconceptinventory.wikispaces.com/
[54] D.E. Denning and P.J. Denning, "Data Security," in *ACM Computing Surveys* **11**(3), 227-249 (1979).
[55] R.J. Anderson, "A Security Policy Model for Clinical Information Systems," in *Proceedings of the IEEE Symposium on Security and Privacy*, 30-43 (1996).
[56] R.J. Anderson, "An Update on the BMA Security Policy," *Personal Medical Information*, 233-250 (1997).
[57] V. Matyás Jr., "Protecting the Identity of Doctors in Drug Prescription Analysis," *Health Informatics Journal* **4**, 205-209 (1998).

[58] P. Doyle, J.I. Lane, J.J.M. Theeuwes, and L.V. Zayatz, "Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies," Elsevier Science (2001).
[59] S. White and R. Chu, "Physics Enrollments in Two-Year Colleges: Results from the 2012 Survey of Physics in Two-Year College", AIP Statistical Research Center, (2013).
[60] P. J. Mulvey and S. Nicholson, "Physics Enrollments Results from the 2008 Survey of Enrollments and Degrees", AIP Statistical Research Center, (2011).
[61] C. Henderson, "The Challenges of Instructional Change under the Best of Circumstances: A Case Study of One College Physics Instructor," *Am. J. Phys.* **73**, 778 (2005).
[62] C. Henderson and M. H. Dancy, "Barriers to the Use of Research-based Instructional Strategies: The Influence of both Individual and Situational Characteristics," *Phys. Rev: ST PER* **3**, 020102 (2007); M. H. Dancy and C. Henderson, "Pedagogical Practices and Instructional Change of Physics Faculty," *Am. J. Phys.* **78**, 1056-1063 (2010).
[63] Association of American Universities Undergraduate Stem Education Initiative: http://www.aau.edu/policy/article.aspx?id=12588 (2011).
[64] C. Henderson and M. Dancy, *Increasing the Impact and Diffusion of STEM Education Innovations*, A White Paper commissioned for the Characterizing the Impact and Diffusion of Engineering Education Innovations Forum, New Orleans, LA, Feb 7-8, 2011. (2011); M. Dancy and C. Henderson, *Barriers and Promises in STEM Reform*, Paper commissioned for Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education, The National Academies Division of Behavioral and Social Sciences and Education. (2008).