Authentic, Career-Specific, Discovery Learning Projects in Introductory Statistics

Project Description

Overview

Many liberal arts, business and education majors require a statistics course, as do STEM majors in health sciences and engineering. Statistics-based research informs pharmaceuticals licensing, political polls and other thorny, real-world issues. Innovative teaching methods in statistics will benefit many citizens, mostly those in non-STEM disciplines. The research literature and a preliminary study indicate the innovations most likely to improve student outcomes will connect course topics to real-life applications.

A prototype for authentic, career-specific, discovery-learning in statistics has been developed at North Georgia College & State University. Project coordinator Dr. Robb Sinn, co-Principal Investigator Dr. Dianna Spence and Faculty Associate Ms. Andrea Hendricks will develop materials based on the proposed model of instruction and scientifically evaluate their utility.

This CCLI Phase 1: Exploratory proposal targets the "Creating Learning Materials and Teaching Strategies" and "Developing Faculty Expertise" components of the cyclical model for improved instructional practice. The materials produced will be a guidebook for instructors and a workbook for students with career-specific, problem-solving units. These materials will allow any instructor to implement this model of teaching and can supplement any standard text in a first statistics course.

This initiative has the **broader impact** of applying to a wide audience of non-majors taking a service STEM course and STEM disciplines that utilize statistics. Preliminary results indicate learning outcomes will be enhanced in two ways: student achievement will increase, specifically transference and retention of knowledge; and student attitudes toward statistics will be improved. The grant-funded work will deliver this unique course design to more than 300 students and provide workshops to more than 50 high school teachers and college professors. Many high schools now offer AP statistics, and most two year colleges have statistics as one of their core offerings. This work benefits students and instructors at all these various levels.

Regarding **intellectual merit**, the proposed methods of instruction are based on widely-accepted best-practice pedagogies such as active learning, collaborative groups, integral use of relevant technologies and authentic, career-specific projects. A working prototype is in place and initial findings show promise. The pilot evaluation will include a valid and reliable survey instrument to compare the achievement and affective student outcomes between a treatment group and control group. Research subgroups will include a four-year university, a two-year college and a high school. The results will be of interest to mathematics education researchers, and the materials will help statistics instructors across the country and at many different educational levels.

Research has shown that effective teaching strategies can improve performance outcomes, attitudes toward mathematics and the transference of problem-solving strategies to non-mathematics and non-school settings. Research and best-practice educational methods informed the development of the prototype teaching framework. These tenets are foundational and a brief review of the literature (see below) summarizes important findings:

- Learning must be active.
- Leaning endures longer when students discover concepts themselves.
- Learning is enhanced by working with collaborative teams and relevant technology.
- Learning is improved by authentic problem-solving that relates to prior interests.

The development of this prototype has been an ongoing curriculum development project at NGCSU. The initial pioneering by PI Sinn (Spring 2004 – present) has been further developed by co-PI Spence who recently utilized the method in two sections of statistics (Spring 2006). The pair have honed a student project structure that guides the discovery-learning process in authentic, real-world settings. Sinn's statistics section this semester worked in career-grouped teams with career-related research questions.

Two investigations give preliminary indications of the prototype's usefulness. A survey indicated the co-PI's three sections (n = 88) had significantly improved attitudes toward the real-world usefulness of statistics compared to a two-section control group (n = 53), and a qualitative textual analysis of project write-ups from Robb's career-specific section (n = 33) indicated approximately half the students strongly connected course material with real-world settings. These findings exemplify the outcomes anticipated based on educational best practice and mathematics education research (see below).

Measurable Objectives

The syllabi of introductory statistics typically cover procedures chosen from seven basic types: regression, confidence intervals, chi-square, ANOVA and the three types of t-test/z-test (one sample, independent samples and dependent samples). By strategically arranging topics, instructors can utilize authentic, discovery learning projects to help students develop deeper insights into the practice of statistics-based research.

The proposed initiative will complete the following three objectives: (1) development of the model and supplementary materials required to implement it including instructor guidebook and student workbook, (2) evaluation of student outcomes in a pilot project utilizing the models and materials, and (3) dissemination of both the supplementary materials and the results of the pilot.

Each objective is measurable. The first objective is a concrete deliverable. The supplementary materials are the linchpin of the proposed research study and have been partially developed. Objective 2 will be measured during the pilot study utilizing a survey instrument in a treatment vs. control group study. Objective 3 will be met by offering the 1½ day professional development workshops to 50 high school teachers and professors, by presentations at professional meetings and publications in peer-reviewed journals, by maintaining a resource website for statistics instructors, and by seeking publication of the materials (see letter of support from W. H. Freeman's statistics editor).

Detailed Project Plan

The majority of college students do not major in a STEM discipline, yet teaching service courses for the liberal arts is possibly the most important (yet least-loved?) duty of mathematicians. The introductory statistics course can be practical, beneficial and interesting to a wide range of students. The proposed model of instruction will help mathematicians structure statistics courses within a research-tested, best-practice format that will improve both performance and attitudinal outcomes. The positive impact will be felt in nearly every discipline across the collegiate educational spectrum and will foster collaborations between mathematicians and many non-STEM disciplines.

The description below will first demonstrate the pedagogical need for the prototype instructional format by briefly reviewing important research literature and two preliminary research investigations. A project timeline and specific description of the work to be conducted follows. A final sub-section details the capability of the research team to conduct the project, the overall evaluation of the project, and dissemination.

Assessment of Need: Literature Review

Development of pedagogically sound statistics instruction is essential. Statistics educators have repeatedly suggested improvements, especially ones that focus on implementing the scientific method utilizing authentic statistical experiences, but these calls for improvement have not been widely heeded (Bryce, 2005). When best-practice pedagogies have been implemented in statistics courses, the results have been positive for achievement and for improved attitudes toward statistics. There is a strong indication that apprentice learning, a modality wherein students complete real-world mathematics in authentic settings, develops better conceptual understanding as well as better transference of knowledge to non-mathematical and non-school settings (Boaler, 1998). Researchers found statistics courses based on more constructivist models improved student attitudes toward statistics and that personal relevance is important for successful learning in statistics (Mvududu, 2003). A researcher used case-study methodology to evaluate a real-world, project-based approach to learning statistics and found that students learned more from the project than from any other instructional component of the course. The researcher further reported improvements in student motivation (Yesilcay, 2000).

There is also evidence connecting positive student outcomes and attitudes toward mathematics as well as motivation for learning mathematics. Researchers performing a meta-analysis of 113 mathematics education studies found a significant influence of attitude toward mathematics upon achievement in mathematics (Ma & Kishor, 1997). Results from an exploratory study using an activity theory model based on Vygotsky's work suggests instructors in statistics courses would do well to consider variables from the affective domain as an integral—not peripheral—part of the statistical learning process (Gordon, 1995). Additionally, interest in mathematics is an increasingly important factor in course selection as students finish secondary school and move into college (Köller, Baumert, & Schnabel, 2001).

A review of the literature on using computer simulations in statistics courses found little current evidence that simulations improve student outcomes (Mills, 2002). Mathematicians may be the most abstract thinkers of all STEM disciplines and are likely entranced by thought experiments or simulations, but students in liberal arts service courses often prefer more concrete activities. Hence, the co-PI's offer the following definition: ***authentic*** statistics activities indicate throughout this document "the collection of real data pursuant to a student-developed hypothesis and connected to a prior interest."

Middleton and Spanias (1999) conducted a review of the literature surrounding motivation to learn mathematics. They reported that careful design of instruction can strongly influence student motivation for mathematics achievement which increases the likelihood students will choose to take future mathematics courses. They provide a beautiful summation of the goals of the proposed initiative: "Students must understand that the mathematics instruction they receive is useful, both in immediate terms and in preparing them to learn more in the fields of mathematics and in areas in which mathematics can be applied (e.g., physics, business, etc.). Use of ill-structured, real-life problem situations in which the use of mathematics facilitates uncovering important and interesting knowledge promotes this understanding" (p.81). The reports cited in the CCLI Program Solicitation offer similar recommendations. The proposed initiative merges well with the relevant research and teaching recommendations for STEM courses.

Assessment of Need: Preliminary Research

A key objective for this model of instruction is to improve affective outcomes. In a treatment-control design, a six-question survey of attitudes toward the usefulness of statistics in real-world settings and in one's future career was used to compare Sinn's and Spence's sections to a control group at NGCSU. The Cronbach alpha coefficient obtained for the instrument was 0.85 confirming its consistency and reliability. An independent samples *t*-test was significant at the 0.10 level and offers preliminary evidence that students who are exposed to this instructional model benefit by connecting course concepts to real-world ideas and their future careers.

	Sample Size *N*	**Sample Mean** $\bar{x}$	**Sample Standard Deviation** *s*
Treatment Group	88	4.38	0.99
Control Group	53	4.13	0.88

(For more details, see supplementary document.)

A qualitative analysis was conducted on the text of the final project write-ups in Robb's section (n = 33). The final paragraph of each student's project report was a reflection guided by the two questions "What did you learn from this class?" and "What did you learn from doing this project?" Each student's paragraph was coded by the research team according to the following scale:

Categories	**IV**	**III**	**II**	**I**
Attributes Used for Coding	Explicit evidence of both real-world and career-specific connections	Explicit evidence of either real-world or career-specific connections	Implicit evidence of either real-world or career-specific connections	No evidence of any real-world or career-specific connections
Number of Students	2	15	5	11

Students in Categories III and IV gave specific evidence of how they had used statistics already. For example, one student wrote that "the concepts taught in class I actually feel I will be able to take with me in the real world" and mentioned having already used course knowledge in another class. One student suggested further interest: "I will probably take another statistics class because I do think it will help me out with real life situations." Two students mentioned how the projects enhance learning: "From this class we have learned that we can apply stats to everyday life to prove or disprove common theories." Students repeatedly mentioned the real-world value of statistics, for example: "I like this class better than other math courses because it is more relevant to real life situations. This project is the perfect example of relating course material to real life."

The qualitative analysis details are available in the supplementary materials along with verbatim reproductions of the final paragraph of every project write-up submitted (Sinn's section, Spring 2006). Half the students provided evidence they realized the important contributions statistical analysis makes to real-world knowledge. The results of both the survey and the text analysis indicate the benefits one would expect from the research literature are likely to be realized during this project.

Project Timeline

The objectives are sequential and form a three-phase implementation plan. The pre-pilot developmental phase (Objective 1) will comprise Summer and Fall 2007.

Project Timeline & Overview

	Tasks	Timeline	Work Product
Phase 1	Materials Development Pre-Pilot Instrumentation	Summer 2007 Fall 2007	Survey Instrument Instructor Guidebook Student Workbook
Phase 2	Pilot Testing	Spring 2008	Data Collection for: • 4-Yr College Implementation • 2-Yr College Implementation • High School Implementation
Phase 3	Dissemination	Summer 2008 Fall 2008	Professional Development Workshops Research Article Submitted Project Website Operational Professional Conference Presentations

The instructor guidebook and student workbook will be developed by co-PI's Robb and Dianna and aided by the NGCSU multidisciplinary team during the Fall semester. During this phase, the survey instrument will be developed with the help of educational research consultant Dr. Frank Pajares of Emory University (see letter of support). NGCSU statistics classes will be used to validate the instrument.

The pilot phase (Objective 2) will take place during Spring 2008. Faculty Associate Hendricks will implement the instructional model in her statistics courses at Georgia Perimeter College (GPC), and an AP Statistics course will be taught using these methods at Forsyth Central High School (see letter of support). The final project phase will disseminate the research findings and developed instructional materials. Research findings will be evaluated by the co-PI's in Summer 2008 and subsequently submitted to a peer-reviewed journal, with the professional development workshops and other presentations occurring in Fall 2008.

Project Implementation Plan

Materials Development. Based on the mathematics education literature, Dr. Sinn and Dr. Spence developed a prototype course structure at NGCSU that allows for authentic research experiences on the part of students in the introductory statistics course. By placing regression topics early in the course, in-depth team projects (with tasks stretched over a six-week timeline) are completed about halfway through the semester. Research questions for the projects are developed by three-person teams based upon common career interests.

The projects require students to think about survey design, representative samples, outliers, scatter-plots, strength of correlation and prediction equations. Students enter their data sets into Excel and use spreadsheet functions to analyze their data both numerically and graphically. The teams present their results to the class and submit a written report of their findings. A numeric summary of each data set generated is then posted on the Internet. Students access "population" statistics for their final group comparison projects that utilize hypothesis testing procedures such as *t*-tests. Projects are team-based, utilize technology, require work products over an extended timeline (a month or more) and include all aspects of study, survey and sampling design, data collection, statistical analysis and reports of findings.

The drawback of the prototype structure is two-fold: assessments of the projects can be both difficult and time-consuming, and the professor has an advisory role in which some expertise in statistics-based research methods may seem necessary. To help ameliorate these issues, the instructor guidebook will contain complete resources for conducting the career-specific, discovery-learning projects in authentic settings. One workbook section will present a "how-to guide" for instructors with practical suggestions for implementation including grading rubrics, suggested timelines, sequence of topics and technology tips. Pre-designed project templates and grading rubrics will help with the assessment difficulty, and the ready-made collection variables and mini-instruments offered in the guidebook will reduce the need for methodological expertise.

Co-PI Spence personally confirms the usefulness of the prototype structure. She implemented the prototype herself for the first time in two sections of statistics this Spring semester (2006). While she has expertise as a mathematics education researcher well acquainted with quantitative research methods, she found the existing components of the instructor guidebook to be a vital and valuable component of her course development. The quality of the materials led directly to a successful implementation of the project structure during her first time teaching the course.

Faculty Associate Hendricks of GPC has extensive collegiate mathematics teaching experience (14 years), but she has limited experience with teaching statistics (1 year at start of pilot) and little experience with statistics-based research. These attributes are an intentional part of the evaluation of the pilot intended to assess the difficulty of implementing this course design without a practitioner's expertise in applied statistics.

The proposed initiative would further develop the prototype structure with the aid of a multidisciplinary statistics team. The team would include participating faculty from seven NGCSU departments who would be given $1,000 consultation incentives for their help in developing project variables and research ideas for the projects. Collaborating departments are of two types. First, departments like Psychology & Sociology, Political Science & Criminal Justice, and Business require the introductory statistics course as a prerequisite course for their own research methods course(s). Second, several majors require the course in preparation for graduate school or implementation of research-based practice, for example, Biology, Nursing, Physical Therapy, and Education.

The work of the multidisciplinary team will provide students with access to many more variables to be studied, especially ones specific to the disciplines mentioned above. Collaborators have already been selected in many of the departments (see letters of support). To say the least, those in other departments have met the idea of statistics instruction relevant to their disciplines with enthusiasm. The expertise of many practitioners will be codified into the supplemental materials, especially the instructor guidebook, so that all instructors can successfully implement these authentic, career-specific, discovery-oriented student activities into their courses.

A second course component would be simultaneously developed by this multidisciplinary team: career-specific problem-solving units from ten different majors with about fifty problems each. The problem-solving units (which form the student workbook) will be a serendipitous outgrowth of the collaboration with the multidisciplinary team and the research necessary to best develop the instructor guidebook. As potential variables and study ideas are catalogued, examples of current statistics-based research will be found that are accessible at the introductory level. The goal is to obtain fifty problems from each of these ten disciplines that are related to typical introductory statistical procedures. The problems will be organized by discipline and would total more than 500 exercises. The workbook can then be used in place of typical course examples to provide specific students with examples that will be of importance to them in their careers and future academic endeavors.

Pilot Phase. The pilot evaluation involves a four-year college (NGCSU), a two-year college (GPC) and a high school (Forsyth Central High School). The major research instrument will be a survey that measures understanding of statistics ideas as well as affective and attitudinal variables, especially the perceived real-world relevance of statistics. To help develop and validate the instrument and verify its reliability, the co-PI's have enlisted the help of Dr. Frank Pajares, professor of education at Emory University (see letter of support). Among his many credentials, the Science National

Honor Society website specifies that Dr. Pajares "serves on six editorial boards, is a Fellow of the American Psychological Association, co-editor of the book series Adolescence and Education, and associate editor of the Journal of Educational Psychology." Dr. Pajares is known to and has worked with both co-PI's in the past and is interested in lending his immense expertise to the evaluation phase of the project.

The research design will compare treatment groups at NGCSU, GPC and Forsyth Central High School to control groups at both the colleges. At NGCSU, the four-year college, both treatment and control groups will comprise approximately six or seven sections (n = 225). At GPC, the two-year college, approximately two sections of statistics taught by Faculty Associate Hendricks will serve as a treatment group (n =50) along with a similar number in a control group. The high school will have a single class in the treatment group (n = 25) and no control group.

When minors are involved in the study, most educational researchers avoid treatment vs. control designs, and school boards often refuse approval for this type of study. Thus, the high school will have a treatment group only. The study will not be compromised. The aggregated AP Statistics exam results for the research district and other Georgia school districts are available in the public domain can be used to assess the conceptual understanding of statistics demonstrated by the high school treatment group. The affective gains in the high school treatment group can be compared to the other active research groups which will serve as a baseline for interpreting the results.

The survey instrument alone would be quite valuable to those interested in improving statistics instruction. Once validated and deemed reliable, the instrument could be used to determine the utility of other innovative methods. Dr. Michael Bodri, NGCSU's Dean of Health and Natural Science and mentor to the co-PI's, proofread this document and pointed out that STEM-specific statistics instruction should not be neglected. With this instrument or another based upon it, educators teaching intermediate and advanced STEM statistics courses could better determine how to help students understand the research significance of statistics within their chosen STEM careers.

Dissemination. Dissemination efforts will be considered successful if four requirements are met. First, two professional development workshops will be conducted for a total of 50 statistics instructors at least half of which are expected to be high school teachers. This goal is supported by an NSF report that noted the introduction of advanced placement statistics in high schools dictates that training and education about teaching statistics must be more vertically focused (Lindsay, Kettering, & Siegmund, 2004). Second, a project website will be developed and maintained. Third, the materials will be submitted for possible print publication. Finally, several conference workshops will be conducted at professional meetings.

The most significant dissemination component is the 1½ day workshop offered to high school statistics teachers and collegiate mathematics faculty. Participants will receive detailed training in using the proposed instructional method and course materials. The workshop will demonstrate how to quickly and fairly assess student presentations and papers, how to utilize the various technologies needed, how to troubleshoot student project designs and questionnaires, and how to integrate courseware and the Internet with the project structure. Full-scale workshops will be hosted at NGCSU in Fall 2008, but condensed workshop formats of two hours and four hours will also be developed. The co-PI's will submit proposals to present these micro-workshops at several NCTM, ICTCM, AMTE and MAA/AMS regional and national meetings.

A website will be created and maintained so that workshop participants and other interested instructors have access to all materials. Statistics textbook publisher W. H. Freeman has agreed to strongly consider these supplementary materials for publication (see letter of support, Ruth Baruth). Beyond the four criteria above, the co-PI's also plan to submit the results of the research study to at least one peer-reviewed research journal. Professional presentations based on the results will also be offered in relevant mathematics education research communities. Opportunities will be sought to disseminate these findings and this course structure beyond introductory statistics into STEM statistics course instruction where results indicate fitness and suitability.

Capability

NGCSU is ideally suited for the collaboration necessary: a small, public, liberal arts college with at least seven departments whose ten or more majors utilize statistics in their undergraduate coursework. The establishment of a multidisciplinary statistics team, the recruitment of top educational research consultant, and the extensive collaboration network which includes both a 2-year college and a high school all provide strong evidence of the capabilities of these co-PI's to implement and complete this project (see letters of support).

Within the Math/CS department, three computer labs provide student access to statistics software including SPSS, Excel and Fathom. A laptop computer lab is available for use during statistics classes. Dr. John Cruthirds, Math/CS chair, has written a letter of support that commits these facilities for grant-related purposes to develop, implement and evaluate the project materials.

Co-PI's Sinn and Spence are mathematics educators whose survey-based research utilizes multivariate statistics and both have research experience studying mathematics achievement variables and variables in the affective domain. Both typically teach two introductory statistics courses and Sinn additionally teaches the upper-level Probability and Statistics sequence for departmental majors. Ms. Hendricks has been teaching mathematics for fourteen years at the collegiate level but will be teaching statistics for the first time in Fall 2006. This level of experience will enhance the results of the pilot offering evidence of the model's utility when attempted by a non-expert in statistics-based research methods. Further, Ms. Hendricks sterling reputation and years of service at GPC will facilitate her role as research coordinator both for the treatment and control groups at GPC. Co-PI's Sinn and Spence have both taught at the high school and college levels, experience that will facilitate the vertical integration components of the project.

Conclusion

This authentic, career-specific, discovery-learning approach to statistics instruction appears likely to produce the quality, relevance and impact common to NSF funded projects. The pedagogy is student-focused in a way that contributes to the STEM education knowledge base. The **broader impact** of the proposal is invigorating: Students in non-STEM majors even as early as high school would be the primary beneficiaries of improved instructional methods. Many STEM majors would benefit by early connection of statistics to their fields of study. The improvement in affective variables would be likely to influence how students think about mathematics by relating their statistics-based investigations to authentic, career-specific inquiry. More than 300 students and 50 instructors will be directly affected by the grant initiative, and immense initial promise has been demonstrated based on the research literature and the preliminary research conducted on the NGCSU prototype.

Development of these materials will foster faculty discussion across many departments, as will their subsequent use. These discussions may introduce collaborations that not only enhance the STEM educational community but also the outreach of STEM educators to faculty in the social-behavioral sciences, business, education and political science. Society will benefit from more sophisticated consumers and practitioners of statistics.

The **intellectual merit** of the proposal includes its framework of well-accepted, best-practice pedagogies such as active learning, collaborative teams, discovery learning, integral use of relevant technologies and authentic and career-specific projects. A prototype has been field-tested with encouraging results. The pilot phase will rigorously test the merits of the proposed model of instruction at several educational levels and in varied formats. The research modality is well-accepted practice in educational studies, and the research team has the capacity to develop and deploy this elegant research design.

In sum, the proposed CCLI development project is well-grounded in mathematics education research and best-practice pedagogy. The implementation of these methods and materials appear likely to improve outcomes for a broad spectrum of students extending throughout and well beyond STEM disciplines. The primary investigators and sponsoring institutions are ideally situated to develop, evaluate and disseminate the proposed teaching method and corresponding materials.