Can research inform us about the efficacy of University STEM education?

Joseph A. Heppert, Associate Vice Chancellor for Research, University of Kansas

Accountability for Student Learning and Its Potential Effects on the Public Research University: Creating more effective models for science, mathematics, engineering and technology (STEM) education represents one of the perennial problem/opportunity scenarios for the future of U.S. public research universities. Trends since the 2007 recession show enrollment in STEM fields on the increase [1], which provides a ready-made response for universities engaged in discussions of their contributions to regional and national economic development. However, improving the retention of students in STEM majors has been a longstanding challenge for universities. Based on recent discussions it appears that many states intend to incorporate student retention, including retention in STEM majors, as a metric in the evaluation of the efficacy of public university performance. [2]

Resulting institutional efforts to improve the quality of STEM instruction and learning outcomes of STEM students will provide an opportunity for public research universities to broaden the scope of the university research mission through the application of rigorous, quantitative social science research methods to our own efforts to drive educational improvement. Employing the methods and metrics of research to studies of undergraduate STEM curriculum change may insulate public research universities from specious arguments that our education and research activities are something other than indivisible aspects of our mission, and from the even more destructive argument, promulgated by some, that the research activities of public research universities detract from our role in teaching undergraduate students. [2]

Over the past 15 years, the increased climate of accountability around the use of taxpayer funding has come to rest on the U.S. public higher education system. Key accountability metrics commonly embraced by both State governments and organizations influencing national higher education policy include both student retention in college and time to degree. [3] These metrics present challenges for traditional models of university STEM instruction, which are perceived to contribute to higher than institutional average attrition from the ranks of STEM majors.

In a ground breaking 1999 study, Seymour and Hewitt examined the motivations of students who leave degree programs in STEM majors, but go on to complete university degrees in alternative academic fields. [4] The study related a number of the factors contributing to this attrition to characteristics of traditional

STEM curricula: Overwhelming amounts of vocabulary, perceptions of poor teaching, and loss of interest in STEM subject matter. These factors relate to some of the challenges facing U.S. Engineering programs, where degree obtention rates of 60 percent represent a national average. [5] Given the current focus on degree completion and time to degree as metrics for university success, these challenges could, if left unaddressed, become a threat to the structure and mission of public research universities.

Research on the nature of human learning has provided a window into effective solutions to these challenges. Many instructional models that accommodate a broader range of learning student learning styles, improve success in learning and increase student engagement with subject matter are based on the cognitive development theories of Piaget. [6] Flipped classroom and peer instructional models which are intended in part as vehicles for enhancing student success and self-efficacy are based on constructivist learning theories. Constructionist instructional models tend to be more student-centered than traditional direct instruction methods. (Interestingly, recent asynchronous instructional methods, including massively open online courses (MOOCS) may, but often do not, employ constructivist-derived learning models.)

While constructionist learning models hold promise for generating improvements in metrics such as time to degree and degree obtention in STEM majors, they are likely to be relatively expensive in terms of supportive infrastructure and faculty opportunity costs. Furthermore, rigorous educational research studies are usually required to unequivocally establish a connection between constructivist based course interventions and improvements in specific, desirable learning outcomes.

In order to demonstrate both to the Academy and to external stakeholders that STEM curriculum changes are actually having the effect we anticipate, we need to consider whether public research universities should systematically turn the tools of rigorous quantitative and qualitative research inward to study the instructional changes being driven, in part, by a culture of increasing accountability. Though this path would increase the cost of implementing curriculum innovation, it would also provide a range of benefits for our institutions, including:

- Development of a broad, new area of multidisciplinary scholarship that is only practicable on a large scale in research universities,
- Creation of a vehicle that allows faculty who have largely shifted their focus to teaching and learning to contribute more broadly to the scholarly life of the university,
- Validation for the university community about the efficacy of the human resources and dollars expended on these efforts, and
- Development of ready-made talking points for accountability discussions with external stakeholders.

Most top-flight public research universities began to take steps to support faculty efforts to create a scholarship of learning as a follow on to the Boyer report [7], and some have developed extensive, nationally recognized expertise in this area. [8] The key questions we must ask about applying these capabilities to studies of STEM curriculum change are: Can rigorous quantitative educational research answer fundamental questions about the efficacy of university curriculum reform, what are anticipated institutional commitments and costs for these studies, and what are reasonable boundaries for the implementation of such programs?

Commitments and Costs for Educational Studies

Among the most daunting challenge of conducting high quality institutional studies of curriculum change is the rigor of designing experiments that will provide meaningful answers to our questions. Collecting a data set that provides adequate statistical power to study all targeted subcategories of learners, maintaining 95 percent confidence limit standards, cleaning and analyzing data sets with large numbers of variables, ensuring the statistical similarity of control and treatment cohorts in an environment of quasi-experimental design (most institutions and faculty are uncomfortable with random assignment studies), and ensuring compliance with human subjects (IRB) requirements all add to the commitment made when undertaking this type of analysis.

A compounding factor associated with such studies is that creating and implementing a relatively straightforward curriculum innovation in a single course, together with designing a course evaluation and collecting and analyzing student outcome data can easily comprise the topic of an entire doctoral thesis. This timeline is problematic for studies of STEM curriculum innovation in research universities where primary interests may lie in longitudinal questions relating the influence of large-scale curriculum change to post-graduate outcomes. Such studies require an extended timeline and more careful research design than studies of a single innovation in an individual course. Achieving this goal would require long-term collaborations among faculty teaching STEM courses, capable quantitative and qualitative educational researchers, and full time university staff dedicated to providing continuity in the study.

Appropriate longitudinal evaluation of curriculum change can be relatively expensive. Guidelines for budgeting evaluation studies in NSF curriculum innovation programs call for commitment of as much as 15 to 20 percent of the total project budget. In larger more comprehensive curriculum innovation projects, this can amount to hundreds of thousands of dollars for a longitudinal evaluation. While quality analyses of student outcomes can be built into university courses for a far lower level of cost, the magnitude of resources required to carry out these studies requires a degree of surety that the study will provide useful outcomes for the institution, as well as careful consideration of the scope and objectives of the study.

Potential Applicability and Utility of Educational Research

First and foremost, we need to ask whether rigorous scientific studies of learning innovations in STEM curricula can provide useful insights into benefits for our students. My own scholarly STEM discipline, Chemistry, is a good context in which to address this question. Traditionally structured university Chemistry curricula have many of the characteristics identified as problematic for student retention in STEM majors in the original work by Seymour and Hewitt: Chemistry courses are built upon abstract concepts, are laden with vocabulary, and require facility with algebra and more advanced mathematics from the outset.

Moreover, over the first three years of study, the Chemistry curriculum swings from algorithmically based material, to subjects requiring substantial memorization, and on to material where calculus becomes the lingua franca. Opportunities to create a synthesis of these different perspectives on the nature of Chemistry often do not occur until the senior year of undergraduate study, or even well into the graduate experience. The initial two years of the undergraduate chemistry sequence have the unfortunate reputation of being gatekeeper courses.

These factors make the Chemistry curriculum a useful test bed for studying whether the application of rigorous educational research methods to the study of new and modified STEM curricula can inform us about improvements in student learning, attitudes and motivation. The three brief examples that follow will illustrate this is possible and that such studies can also yield interesting, unexpected insights.

In my group Dr. Danielle Barker recently pursued a study of whether asynchronous mathematics learning tools built using a constructionist educational framework would improve student performance and self-efficacy in a freshman chemistry course for science majors. [9] Facility with algebra and algebraic reasoning are among the most critical skills required for success in freshman chemistry courses; consequently, these are topics that tend to be emphasized in the initial weeks of Chemistry instruction.

The fact that standardized examinations developed by the American Chemical Society indicate this area to be a weakness in up to 30 percent of our students has tended to reinforce the early coverage of chemistry-related algebra concepts. Unfortunately, subjects such as significant figures, ratios, and negative logarithmic scales are scarcely the most charismatic and integrative aspects of the discipline of Chemistry. Dr. Barker's study was intended to determine whether these subjects could be covered asynchronously, and whether this change might enhance student achievement and self-efficacy in the course. Students participated in a series of 40 online chemistry oriented mathematics tutorials over the course of the first semester.

Studies of student achievement showed that benefits were dependent on student persistence though the majority of the units. Students persisting through the tutorials showed nearly a full grade point improvement over a control group and half grade point improvement over students receiving traditional in-class math concept instruction. (Note: Students who persisted through 35 or more units started the program with demographic and academic characteristics that were indistinguishable from the class as a whole.) Students completing the tutorial also showed sustained higher levels of self-efficacy with respect to chemistry content knowledge than their peers.

Ms. Linda Myers is currently concluding a study of whether peer-led undergraduate supplements (PLUS), group work problem-solving assignments coordinated by a trained student leader, improve student achievement in freshman chemistry. These learning tools are related to peer-led team learning (PLTL) and process-oriented guided inquiry learning (POGIL) strategies that have been successfully employed as active learning supplements to lecture and laboratory experiences in other contexts. [10, 11] Ms. Myers' studies show that students persisting in weekly PLUS session show a 14.5 percent improvement on chemistry examinations over comparable peers. This result is moderated by gender, with male students experiencing higher benefits than female students. Interestingly, the cohort of pre-pharmacy students in the course showed no overall benefit from participation in PLUS sessions. Based on prior academic performance, pre-pharmacy student can be categorized as being among the most academically capable of the major-cohorts in freshman chemistry.

Finally, Dr. Deblina Pakhira has examined whether the common practice of allowing students to choose whether to enroll concurrently in Organic Chemistry lectures and laboratories has any effect on student learning and achievement in these classes. [12] This project relates to an ongoing study of whether practicing components of the discipline of chemistry within the laboratory benefits student learning.

We were surprised to discover that students choosing to enroll concurrently for the lecture and laboratory, and students who enroll first in the lecture course and then enroll in the laboratory in a subsequent semester begin Organic Chemistry with statistically indistinguishable demographic, academic, and motivational characteristics. Despite these initial similarities, Dr. Pakhira's 3-year study demonstrated students choosing concurrent enrollment during the first semester Organic Chemistry course showed a quarter grade point average advantage in achievement over their colleagues enrolled in only the lecture portion of the course. During the second semester course, this advantage increased to a half grade point for students choosing concurrent enrollment. Moreover, students concurrently enrolled in lecture and laboratory showed higher longitudinal motivation and self-efficacy regarding the Organic Chemistry course sequence.

Recommended Boundaries for the Evaluation of Curriculum Reform Efforts

As these examples show, it is possible to gain useful, sometimes surprising insights from rigorous evaluation of curriculum reform efforts. However, the resources required to conduct such studies on a large scale should lead us to engage in a careful consideration of the circumstances that justify an intensive evaluation of STEM curriculum innovation. The following list identifies some of the characteristics that might reasonably trigger a need for institutionally supported longitudinal STEM curriculum studies:

- True novelty in curriculum design, instructional practice or application of technology,
- Networks of STEM curriculum change across multiple courses or disciplines that together are intended to create a broader impact on student learning and outcomes,
- Curriculum interventions that require substantial investment of institutional resources,
- Experiments in curriculum change that are part of a broader, national reform study,
- Changes that may have high stakes consequences for students, faculty and instructors, and the institution, and
- Efforts to engender longitudinal (post-graduation) advancements in student knowledge, skills and abilities.

Finally, we need to consider the range of questions we should strive to address through institutionally supported research studies. The following are examples of big picture questions that should drive our curiosity in this area:

- At what threshold of curriculum change (class component, course, core education program, major curriculum) do we observe the onset of specific desired benefits in student learning and success in degree programs?
- Have curriculum innovations created noteworthy enhancements in student ability to obtain and apply new knowledge, skills and abilities in future professional endeavors?
- Beyond performance in individual courses, what are the key educational metrics that we want to promote...increased retention in majors, reduced time to degree, improved rate of degree attainment, etc.?
- How have curriculum changes influenced retrospective student perception of educational value?
- What degree of improvement in student learning, perception and attitude is sufficient to justify a specific level of institutional investment in curriculum innovation?
- Is the institutional commitment to longitudinal research on STEM curriculum innovation contributing to expanding the productivity of STEM and education faculty researchers?

Given the stresses that external accountability is exerting on university budgets and faculty researchers, it is to our advantage to demonstrate that resources aimed a STEM curriculum enhancement are providing the anticipated benefits for our students. We have the tools of research at our disposal, faculty who could benefit professionally from partnering in such studies, and the need to move away from an anecdotal narrative for evaluating the efficacy of educational change. This process can contribute to protecting the diverse, interrelated missions of public research universities and provide a narrative for engaging a sometimes-skeptical public in the discussion that the research and educational missions of the university are indivisibly linked.

Acknowledgements.

I wish to thank the KU Office of Diversity Training; the KU Office of Institutional Research and Planning (OIRP); and Drs. M. Danielle Barker and Deblina

Pakhira, and Ms. Linda Myers whose thesis research is highlighted in this work. I also wish to thank Ms. JoAnn Williams (OIRP) for her work with obtaining KU institutional data. Portions of this work were supported by the National Institutes of Health (NIH IMSD R25 GM062232-08).

References

1) Scott Jaschik, 2014, "The STEM Enrollment Boom." Inside Higher Education.

2) Patricia Kilday Hart, 2011, "Perry's vision for colleges has professors seeing red." The Chronicle of Higher Education.

3) Kansas Board of Regents, 2011, Foresight 2020 Strategic Agenda

4) Elaine Seymour and Nancy Hewitt, 1999, "Talking About Leaving: Why Undergraduates Leave the Sciences." Westview Publishers.

5) John Marcus, 2012, "High Dropout Rates Prompt Engineering Schools to Change Approach." http://blogs.ptc.com/2012/08/06/high-dropout-rates-prompt-engineering-schools-to-change-approach/

6) Hildegard, Bauer, 1975, "Theories of Learning", Ch. 10, Piaget's Developmental Psychology.

7) Carnegie Foundation for the Advancement of Teaching, 1995, Reinventing Undergraduate Education: A Blueprint for America's Research Universities."

8) For example, see University of Colorado, Center for STEM Learning, http://www.colorado.edu/csl/

9) Danielle Barker, 2011, "Constructivist-Based Asynchronous Tutorial to Improve Transfer between Math and Chemistry Domains: Design, Implementation, and Analysis of the Impact of ReMATCH on General Chemistry Course Performance and Confidence", Doctoral Thesis, University of Kansas, Lawrence, KS

10) Pratiba Varma-Nelson and Brian Coppola, 2004, Team Learning in The Chemists' Guide to Effective Instruction, Prentice Hall, V.1.

11) J. N. Spencer, 2006, "New Approaches to Chemistry Teaching", J. Chem. Ed., 83, 528-535.

12) Deblina Pakhira, 2012, "Analysis of the Effect of Concurrent Enrollment in Organic Chemistry Lecture and Laboratory Instruction on Student Learning and Motivation towards Learning Chemistry", Doctoral Thesis, University of Kansas, Lawrence, KS