# Undergraduate Research Experience For STEM Students: Efforts And Outcomes

Chuanlei Zhang, Philander Smith College, USA
Samar Swaid, Philander Smith College, USA

**ABSTRACT**

*Undergraduate research for STEM students involves students who are attending college or universities pursuing a bachelor's degree, majoring in fields related to Science, Technology, Engineering and Mathematics(STEM). Research experience for STEM undergraduates has been viewed as a positive experience that has several benefits such as developing intellectual mentality, enhancing teamwork skills and improving preparation for graduate study. In this paper, we provide an overview of HBCU-UP research experience with STEM students at Philander Smith College for the last few years. Projects in Bioinformatics and Human-Computer Interaction of two Computer Science faculty and their teams are shared. Efforts in implementing research experiences and outcomes of these efforts are discussed.*

**Keywords**: STEM; Computer Science; Undergraduate Research

**INTRODUCTION**

As a process of gathering data, provoking curiosity for pattern recognition and exploration of the nature of data and possibly identifying hidden message, research has been widely conducted in various fields. Undergraduate research involves students who are attending college or universities pursuing a bachelor's degree; and STEM student research activities involved the students who are majored in fields related to Science, Technology, Engineering, and Mathematics. As stated in a report of National Science Foundation, "It is clear that the academic community regards the involvement of undergraduate student majors in meaningful research … with faculty members as one of the most powerful of instructional tools" (NSF,1989, p.6). Four characteristics, "mentorship, originality, acceptability, and dissemination," have been pointed out by Hakim (Hakim, 1998) that can be used to describe undergraduate research experience.

Engaging students in research projects is frequently cited as an effective way to link faculty research and undergraduate teaching (Prince, Felder, & Brent, 2007). Examples can be the involvement of undergraduates in engineering fields (Sabatini, 1997), NSF REU program for Undergraduates (NSF REU), results in the literature survey by Pascarella and Terenzini (Terenzini, Pascarelle, & Bliming, 1996). Besides, such experience shows 'enhance intellectual skills such as inquiry and analysis, reading and understanding primary literature, communication, and teamwork … These benefits are an advantage in any career path' (Lopatto, 2010). 'Faculty mentors agree that undergraduates gain significant education benefits from research experience, that includes increased technical skill, ability to act independently, insight into graduate study and career possibilities, etc.,' according to Ward, Bennet, & Bauer (2002) in their efforts to examine the educational effectiveness of the undergraduate research experience. According to the intensive survey by Pascarella and Terenzini (Terenzini et al., 1996), it shows 'positive outcomes for students who participate in undergraduate research programs, including greater retention in the curriculum and greater likelihood of enrolling in graduate school.' Another study by Hippel (Hippel, Lerner, Gregerman, Nagda, & Jonides, 1998) shows that 'participation in the Undergraduate Research Opportunity Program increased retention rates for some students. In general, this effect was strongest for African American students and for sophomores'. Surveys were conducted in Russell's study (Russell, Hancock, & McCullough, 2007) which suggests that enthusiasm is key to fuel interest in STEM research and higher degrees as well as an early explosion to such experience.

## OVERVIEW OF THE HBCU-UP GRANT RESEARCH AT PHILANDER SMITH COLLEGE

Philander Smith College of Little Rock, Arkansas, is the oldest minority serving, four-year liberal arts college with a total enrollment of more than 500 students. The key division of the College is the Division of Natural and Physical Sciences which serves more than 200 students majoring in disciplines of biology, chemistry, computer science, and mathematics. The division has an articulation agreement with the University of Arkansas, Fayetteville (UARK), to implement the 3/2 Engineering program to allow students receiving an additional engineering degree from UARK. The mission of the Division of Natural and Physical Sciences is " to graduate Science, Technology, Engineering and Mathematics (STEM) students who are academically accomplished and equipped with: comprehensive knowledge in their science fields, modes of inquiry, quantitative reasoning, scientific communication, ethical values that guide their practice, and problem solving and decision making. Students are trained, as researchers, critical thinkers, social justice advocates, leaders, and policymakers, to succeed in post STEM-related graduate schools, to compete in the STEM workplace, and to contribute to educational and economic advances in local, regional, national and international science and technology initiatives". (Philander Smith College Catalouge, 2015, p.3)

On 2012, the Division has been awarded a grant to implement a multi-phase project to improve students' performance and retention in Science, Technology, Engineering and Mathematics (STEM) fields via three arms: (i) improve STEM students preparedness to college life via a summer program called STEPP; (ii) establish a deep understanding of core concepts of STEM introductory courses via the application of computational thinking and cyberinfrastructure resources (Swaid, 2013a, 2015a, 2015b) and (iii) iv) conduct a year-long undergraduate research to enhance students research skills, strengthen preparedness to graduate studies and accelerate student learning and motivation by the involvement in disciplinary research.

This paper focuses on our efforts in undergraduate research with STEM students. As is noted in education research, undergraduate students' participating in hands-on research is widely recognized as a strategy to encourage students to pursue advanced degrees and careers of STEM (Hunter, Laursen, & Seymour, 2006). Inspiring a culture of research among undergraduates improves students confidence and self-esteem (Russell et al., 2007). The Division of Natural and Physical Sciences developed a multi-module interdisciplinary research modules that emphasize inquiry-based activities. The research modules included areas of High-Performance Computing (HPC), Human-Computer Interaction (HCI), Molecular Modeling, Bioinformatics (BI) and Analytics and Optimization. Every year, research faculty would develop a proposal for their research projects. Proposals are shared with students who are provided with instructions on the application process. Considering criteria of students GPA, major, course work, research interest and commitment degree, each supervisor selects his/her group of students. This paper focuses on efforts to engage students in undergraduate research in areas of Bioinformatics of the first authors and Human-Computer Interaction module of the second author.

### Introduction to the Bioinformatic (BI) Project

With the vast availability of biomedical research data, data mining and machine learning algorithms that are designed for pattern recognition have been widely applied on them to understand the nature of the data and interpret them, especially in bioinformatics. Classification, a widely studied topic in data mining and machine learning, is an action or a process of classifying elements into different classes based on shared qualities or characteristics. Binary classification is the task of classifying the elements of a given set into two groups by a classification rule. Different algorithms have been studied and applied to various biomedical datasets, e.g. genomic data, cancer data. In this module, the focus was on using biomedical data from Parkinson's Disease (PD) with a data mining application.

Parkinson's disease is the disorder of the brain that progressively deteriorates the motor function which leads to one to constantly shake (tremors) causing one to have difficulty walking, in their movement and coordination (Parkinson Disease, 2014). In severe cases, Parkinson's affects the voice and hearing as well. Lee Silverman Voice Treatment, also known as LSVT, is one of the most widely practiced treatments for speech disorders associated with PD. It gives feedback after a series of tests on the status of one's voice.

This team, including one CS faculty, one biology student, one math student and two CS students, has been working on such a data mining problem in bioinformatics - application of a machine learning algorithm on the classification

of a biomedical dataset. The project was titled 'Assessment of Voice Rehabilitation Treatment using Selective-voting in convex hull ensembles algorithm.' The research goal here is to gain insight into objective automatic assessment of rehabilitative speech treatment in Parkinson's disease using machine learning algorithm (Tsanas, Little, & Fox, 2014). To be specific, to assess whether voice rehabilitation treatment leads to phonations considered 'acceptable' or 'unacceptable' as the two classes in this classification problem, using the 'selective-voting in convex hull ensembles algorithm' (Kodell, Zhang, Siegel, & Nagarajan, 2012). The data that the team used contains 126 samples from 14 participants PD with 309 features. The abstract was presented at 4th AR Capitol STEM meeting on February 11, 2015, and the STEM Research Forum at Philander Smith College on April 16, 2015.

*BI Phase I - Abstraction*

Beginning at the project, topics on data mining, classification, and related concepts were introduced with a problem-based learning approach. Sample questions were given to students, e.g. defining data mining and classification in data mining, giving examples of classification problems, binary classification problems and discuss how to access classification accuracy. These questions arm the students with the understanding of the basics on classification, its general factors and how to access classification performance.

*BI Phase II - Introduction to Software R*

During the second phase of the project, since the algorithm was originally written in open source software R (R 2013), this software was introduced to students at this stage. Due to the mixed nature of students' major background, they were working as a team in a way that CS students contributed more on the coding assignment while biology students helped others with the understanding of the background of the dataset and the disease.

*BI Phase III - Algorithm Application and Evaluation*

After the introduction to binary classification and software R, students were asked to apply the algorithm to the LSVT dataset. With the help of the faculty, final results of running the algorithm were obtained. The students were asked to evaluate the performance of this binary classification task statistically. Several important indices that are widely used in research were adopted here, that is accuracy - proportion of correct predictions among all samples, sensitivity - proportion of correct predictions among true positives, specificity - proportion of correct predictions among false positives, positive predictive value - proportion of correct predictions among positive predictions and negative predictive value - proportion of correct predictions among negative predictions. Finally, students were asked to make observations on these indices and discuss the average performance for each class for this classification problem.

## INTRODUCTION TO HUMAN-COMPUTER INTERACTION (HCI) PROJECTS

As computing technologies have become integrated into everyday lives, it is imperative to understand aspects of making these technologies user-friendly and accessible to all citizens. As such, undergraduate research studies were designed to introduce theories of human-computer interaction and human-centered computing to understand issues related to design factors, usability, accessibility and technology impact.

**HCI Research Phase II: Technology Use by Elderly and Usability Issues**

Research has found that citizens aged 65 and older are 43% slower at using technology than others. Although, There are improvements over previous designs, still technology designs need to change to better accommodate aging users (Nielsen, 2013). Research was designed to explore factors that might impact elderly from using technologies in their daily tasks. There were three students, two females, and one male, engaged in the study. The research group mainly focused on issues related to the use of computer software such as MS Office, email systems, web browsers and social networking sites. First, a thorough literature review of elderly--related use of technology was applied. Next, students conducted a content analysis of scholarly work to identify barriers to technology use to elderly citizens. Third, the students were asked to develop a plan of actions that would help increase elderly citizens use of technology and in general and computers and the web in particular. Finally, the students implemented the

plan of actions under the supervision of the second author.  Based on literature review, students identified that some age-related changes such as sensory changes, motor changes, and social changes might explain the difficulty in using technology.  The students were surprised to find that old person might not using technology as other adults due to lack of experience (Smith & Swaid, 2014).  Therefore, the students developed a plan of actions in which they would train a group of senior citizens on using computers, software, email system, and social networking sites. The outcomes of this research were as follows:

*Computer Illiterate 2 Computer Savvy(CI2CS) Initiative*

A three-week tutoring sessions at the Patrick Henry Hays Senior Citizens Center, of North Little Rock, AR, were held where a total of seven seniors was trained.  The Patrick Henry Hays Senior Citizens Center is a public facility constructed for the primary purpose of providing senior activities and classes.  The Hays Center is open to persons 50 years and older for a nominal fee.  Training sessions were offered two times a week to teach seniors topics of understanding computer hardware parts, log-on to computes, creating folders, sending and receiving emails, and developing MS Word documents.  The elderly expressed enthusiasm and more confident using computers.  Reviews and assessments illustrated the high retention of information (Smith & Swaid, 2014).

*Philander Smith College's Basic Skills Computer Workshop*

On-site six-weekend training session was held at Philander Smith College in the form of weekend course.  This was the first professorial project ever committed at Philander Smith College to train elderly on using computers and software. Total of two male seniors were registered in the training.  Senior students participated in the training sessions showed an exponential growth in their computer literacy during each session.  These sessions trained elderly on fundamentals of using technology and allowed undergraduate research students to understand the barriers to technology use to some populations. Undergraduate students presented their work in a national conference (Smith & Swaid, 2014; Dehart & Swaid 2016) and were awarded a scholarship to present their research and outcomes in the Clinton Global Initiative University (CGIU) of 2014.

**HCI Research Phase II: Website Quality and Usability Measures**

Regardless of the context of use, website quality is considered one of the key dimensions that may explain usage, satisfaction, and loyalty (Swaid & Wigand, 2007; Swaid & Wigand, 2009).  In this phase, students were engaged in understanding quality and usability concepts of websites and its effect on user's satisfaction. One male student majoring in Mathematics participated in exploring website quality of Open Educational Resources (OER) websites. The study started with understanding the architecture of OER and then identifying quality factors of OER websites. This study applied theories from human-computer interaction to analyze the website architecture of the open education resources and proposed a framework to quantitatively evaluate the website architecture.  Based on data collected from 135 subjects, and a factor analysis, the study revealed that variables of content, navigation, and interactivity are the key factors to use when evaluating the quality of OER websites (Mayfield & Swaid, 2014). The second study focused on identifying quality dimensions of e-commerce websites that sell digital products (e.g., websites themes, software, plugin, etc). A structured content analysis of the top eighteen websites that sell website themes, plugins and software was applied.  Content analysis study revealed factors of (i) visual appeal; (ii) structure; (iii) creativity; (iv) recommendations and (v) assurance to be the main factors to measure e-commerce websites quality.  Research work was published in the 2014 Emerging Researchers National (ERN), 2016 ERN and the 2014 Cyberinfrastructure Day of Philander Smith College.

**HCI Research Phase III: Usable Security**

One of the most important areas in usable security is user authentication (Payne & Edwards, 2008).  Particularly, researchers are interested in user's authenticate and password use.  Usable security present an "innate tension" between usability and security.  Two research studies were designed to explore issues related to usable security. One study focused on the usability of Completely Automated Public Turing Test to Tell Computers and Humans Apart (CAPTCHA), while the second study explored the links between password security policies and passwords strength. A total of four students, three female Computer Science students, and one male Computer Science student,

**216**

were interested in the first study focusing on CAPTCHA usability.  CAPTCHA is a security mechanism that is used by websites to block spam and hacking.  Although a large number of sophisticated CAPTCHAs are present, most of them are unusable.  The study developed usability measures of text-based CAPTCHA by applying "netnography" method of a content analysis, focus-groups and sorting task.  The study not only identified the manifest measures of text-based CAPTCHA usability but also developed the latent symbolic meaning of usability dimensions of CAPTCHA (Swaid, 2013b, Swaid, 2013c, Smith et al., 2013b, Hawkins, Petty, & Swaid, 2014).  The study's findings were published in the 2014 ERN and at the International Conference on Learning and Administration in Higher Education (Smith & Swaid, 2014).

The second project studies the quality of password construction policies and its effect on password strength.  In this study a total of five students, two female biology students, one female computer science student and two male computer science students, collaborated to investigate password policies and its effect on password strength in different websites. A total of 100 websites were identified to analyze their password construction policies considering the entropy of passwords generated.  Research work revealed interesting patterns of password policy, password strength and website type. Study findings were published at the 2016 ERN (Petty & Swaid, 2016) and the 2016 PSC STEM Research Forum of Philander Smith.

## DISCUSSION

This paper has presented the research experience efforts for STEM undergraduate students at Philander Smith College during the last few years. Throughout the projects, students have been exposed to different research topics, as well as various means of presentation of the results, such as conference poster presentation, discussion forums, and paper publication. Such experience shows that students have benefited to 1) develop and build stronger interest in understanding the nature of research, 2) gain confidence when presenting research results in public, 3) develop team collaboration skills, 4) improve communication skills, and 5) improve understanding and be better prepared for their graduate study.

## ACKNOWLEDGEMENT

## AUTHOR BIOGRAPHIES

**Chuanlei Zhang, Ph.D.** is currently an assistant professor at the department of Mathematics and Computer Science Philander Smith College. She received her Ph.D. from University of Arkansas at Little Rock. She had post-doc at Department of Biostatistics of University of Arkansas for Medical Sciences prior to joining Philander Smith College. Her research area includes data mining, machine learning, and artificial intelligence.

**Dr. Samar Swaid** is the Chairperson for the Division of Natural and Physical Sciences and Associate Professor of Computer Science at Philander Smith College of Little Rock, Arkansas. She was awarded her PhD degree in Applied Computing from University of Arkansas at Little Rock on 2007. She is the PI and Co-PI of a number of projects that focus on improving STEM education and broadening participation in STEM fields. Her research spans areas of STEM undergraduate education, computational thinking, cyber-infrastructure use in higher education, human-computer interaction and human-centered computing. Her research is published in peer-reviewed journals, chapter books and conference proceedings.

## REFERENCES

DeHart, J. & Swaid, S. (2016). Developing a Framework to Assess the Quality of Websites Selling Digital Products. 2016. Emerging Researchers National (ERN) Conference, Washington, D.C., Feb 22-24.

Hakim, T. (1998). Soft assessment of undergraduate research: Reactions and student perspectives. Council on Undergraduate Research Quarterly,18, 189-192.

Hawkins, D., Petty, B. & Swaid, S. (2014).  A Comparison between Still-Image CAPTCHA and Video CAPTCHA: A Usability Perspective. Emerging Researchers National (ERN) Conference, Washington, D.C., Feb 22-24.

Hippel, W. V., Lerner, J. S., Gregerman, S. R., Nagda, B. A., & Jonides, J. (1998). Undergraduate student-faculty research partnerships affect student retention. *The Review of Higher Education*, *22*(1), 55-72.

Hunter, A., Laursen, S.  & Seymour, E. (2006, February 10 – 11). Benefits of participating in undergraduate research in science: A comparative analysis of student and faculty perceptions. Paper presented at the Texas Tech University, Lubbock, TX.

Kodell, R. L., Zhang, C., Siegel, E. R., & Nagarajan, R. (2012). Selective voting in convex-hull ensembles improves classification accuracy. *Artificial intelligence in medicine*, *54*(3), 171-179.

Lopatto, D. (2010). Undergraduate research as a high-impact student experience. *Peer Review, 12*(2), 27-30.

Mayfield, J. & Swaid, S. (2014). Quantitative Evaluation of Open Educational Resources: The Website Architecture Approach. Emerging Researchers National (ERN) Conference, Washington, D.C., Feb 22-24.

National Science Foundation (NSF). (1989). Report on the National Science Foundation disciplinary workshops on undergraduate education. Washington, DC

Nielsen, J. (2013). Usability for Web Users. Retrieved from:  https://www.nngroup.com/articles/usability-for-senior-citizens/, on March 16, 2016.

Parkinson Disease*,* PubMed, Retrieved March 22nd, (2016), from PubMed.gov: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001762/#adam_000755.disease.causes.

Payne, P.  & Edwards, K. (2008). A Brief Introduction to Usable Security. IEEE Computer Security. Retrieved from: http://www.cc.gatech.edu/~keith/pubs/ieee-intro-usable-security.pdf, on March 20, 2016.

Petty, B., & Swaid, S. (2016). Measuring the Entropy of Passwords. The Website Architecture Approach. Emerging Researchers National (ERN) Conference, Washington, D.C., Feb 22-24.

Philander Smith College Catalogue 2015-2017, retrived from http://www.philander.edu/docs/Campus-Catalogue-24.pdf on March 22nd, 2016.

Prince, M. J., Felder, R. M., & Brent, R. (2007). Does faculty research improve undergraduate teaching? An analysis of existing and potential synergies. *Journal of Engineering Education*, *96*(4), 283-294.

Russell S., Hancock M., McCullough J.  (2007). The pipeline. Benefits of undergraduate research experiences. Science 316(5824): 548-549.

Sabatini, David A. (1997). Teaching and research synergism: The undergraduate research experience." *Journal of Professional Issues in Engineering Education and Practice 123*.3, 98-102.

Smith, C. & Swaid, S. (2014). The Great American Gap: Elderly and Technology. International Conference on Learning and Administration in Higher Education, May 21st-23rd 2014, Nashville, TN.

Swaid, S. (2013a). Cyberinfrastructure for Undergraduate STEM Education. International Conference on Learning and Administration in Higher Education, *Journal of Learning in Higher Education*. *Fall-2013*, 61-64.

Swaid, S. (2013b). Usability Measures of Text-based CAPTCHA: Application of Think-Aloud Protocol.  International Journal of the Academic Business World. Fall-2013, 63-66.

Swaid, S. (2013c). Scale Development to Measure the Usability of Text-Based CAPTCHA. International HCI 2013, Las Vegas, NV, July 22-26, 2013.

Swaid, S. (2015a). Computational Thinking in STEM Education: The Application of Community of Practice. The Clute International Education Conference, New York, August 2- August 7, 2015.

Swaid, S. (2015b).  Bringing Computational Thinking to STEM Education. AHFE 2015 International Conference, July 25- July 30, Las Vegas, NV.

Swaid, S. & Wigand R. (2009). Measuring the quality of e-service: Scale development and initial validation. *Journal of Electronic Commerce Research*, *10*(1): 13-28.

Swaid, S. & Wigand, R. (2007). Key Dimensions of E-commerce Service Quality and Its Relationships to Satisfaction and Loyalty, proceedings of *20th Bled conference,* Bled, Slovenia.

Terenzini, P.T., Pascarella, E.T. & Blimling, G.S., (1996). Students' out-of-class experiences and their influence on learning and cognitive development: A literature review. *Journal of College Student Development*.

Tsanas, M.A. Little, C. Fox, L.O. (2014). Ramig: "Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 22, 181-190.

Ward C, Bennett J, Bauer K (2002) Content analysis of undergraduate research student evaluations. Retrieved Oct 2016 from http://www.udel.edu/RAIRE