Equal Outcomes 4 All: A Study of Student Learning in ECS

Steven McGee
Randi McGee-Tekula
Jennifer Duck
Catherine McGee
The Learning Partnership
mcgee@lponline.net

Lucia Dettori
DePaul University
lucia@cdm.depaul.edu

Ronald I. Greenberg
Loyola University
rig@cs.luc.edu

Eric Snow
Daisy Rutstein
SRI International
eric.snow@sri.com

Dale Reed
University of Illinois Chicago
reed@uic.edu

Brenda Wilkerson
AnitaB.org
brendadw@anitab.org

Don Yanek
Andrew M. Rasmussen
Chicago Public Schools
DGYanek@cps.edu

Dennis Brylow
Marquette University
brylow@mscs.mu.edu

ABSTRACT

This study investigated patterns in the development of computational thinking practices in the context of the Exploring Computer Science (ECS) program, a high school introductory CS course and professional development program designed to foster deep engagement through equitable inquiry around CS concepts. Past research indicates that the personal relevance of the ECS experience influences students' expectancy-value towards computer science. Expectancy-value is a construct that is predictive of career choices. We extended our research to examine whether expectancy-value influences the development of computational thinking practices. This study took place in the context of two ECS implementation projects across two states. Twenty teachers, who implemented ECS in 2016–17, participated in the research. There were 906 students who completed beginning and end of year surveys and assessments. The surveys included demographic questions, a validated expectancy-value scale, and questions about students' course experiences. The assessments were developed and validated by SRI International as a companion to the ECS course. Overall, student performance statistically increased from pretest to posttest with effect size of 0.74. There were no statistically significant differences in performance by gender or race/ethnicity. These results are consistent with earlier findings that a personally relevant course experience positively influences students' expectancy for success. These results expanded on prior research by indicating that students' expectancy-value for computer science positively influenced student learning.

CCS CONCEPTS

• **Social and professional topics→Computational thinking**; *Model curricula; Student assessment*; *K-12 education*

SIGCSE '18, February 21–24, 2018, Baltimore , MD, USA

ACM ISBN 978-1-4503-5103-4/18/02.
https://doi.org/10.1145/3159450.3159529

KEYWORDS

high school computer science; computer science assessment; computer science attitudes; Exploring Computer Science; Expectancy-Value-Cost; computer science teaching practices

ACM Reference format:
S. McGee, R. McGee-Tekula, J. Duck, C. McGee, L. Dettori, R.I. Greenberg, E. Snow, D. Rutstein, D. Reed, B. Wilkerson, D. Yanek, A.M. Rasmussen, and D. Brylow. 2018. Equal Outcomes 4 All: A Study of Student Learning in ECS. In *SIGCSE '18: 49th ACM Technical Symposium on Computer Science Education, Feb. 21–24, 2018, Baltimore, MD, USA.* ACM, NY, NY, USA, 6 pages. https://doi.org/10.1145/3159450.3159529

1 INTRODUCTION

The Exploring Computer Science (ECS) curriculum and professional development program was developed at the University of California, Los Angeles, with the goal of contributing to broadened participation of women and minorities and increased equity in the field of computer science [18]. Specifically, the ECS curriculum seeks to accomplish this goal of broadening participation by introducing the field of computer science and computational practices in a way that makes the field relevant, engaging, and stimulating for a diverse population of students. The ECS curriculum is composed of activities that are designed to engage students in computer science inquiry around meaningful problems; the ECS professional development program is designed to prepare teachers to implement these inquiry-based activities while also guiding teachers in building a classroom culture that's culturally relevant and inclusive of all students. Prior studies have successfully documented the impact of this professional development on the quality of ECS implementation [13]. Prior studies have also shown that students' perceptions of the relevance of the ECS course experiences influence students' attitudes towards computer science and influence the likelihood that students will pursue further computer science coursework [6,19].

With continued support from the United States National Science Foundation (NSF), a variety of university- and community-based organizations are adopting the ECS program and rapidly expanding its reach to cities across the United States. This study took place during the 2016–17 school year in the context of ECS implementation projects in the state of Wisconsin and in the Chicago Public Schools (CPS). In Wisconsin, there were forty-four teachers who implemented ECS throughout the state. In Chicago, CPS had recently instituted CS as a high school graduation requirement. Over 100 teachers implemented ECS, which fulfills the graduation requirement. In this research, we seek to extend our prior research to investigate the extent to which students' attitudes towards computer science influence the development of computational thinking practices.

2 THE KEY COMPONENTS OF THE ECS CURRICULUM

Key to the design of the ECS curriculum is deep engagement within a community of practice. When computer science is not taught for deep engagement but rather as an abstract academic subject, it privileges access to computer science to mostly Caucasian, male students [16]. To play an integral role in such classrooms, students must master abstract programming for programming's sake. Typically, computer science courses at both high school and college levels have been taught in this abstract way [17]. For non-Caucasian students in low-income neighborhoods, computer instruction has tended to focus on computer applications and has lacked opportunities for engaging in collaborative inquiry [16].

The ECS curriculum is designed to engender deep engagement with important computer science concepts by mimicking important features of communities in which youths participate outside the classroom. General technology use outside of school by youths of all races and genders tends to revolve around making social connections and working on practical problems [15]. Reorienting computer science instruction to be culturally relevant and focused on problem-solving experiences that are meaningful to students has the potential to increase access to computer science content, provide students with integral roles, and create opportunities for students to express themselves [4]. At the college level, computer scientists at Carnegie Mellon made progress at increasing the representation of women in their computer science program by making such changes to the nature of instruction in their introductory courses. Students develop technical fluency through solving problems of interest [17].

At the core of ECS are a set of high-leverage teaching practices [14] that support the three interwoven teaching strands of ECS: equity, inquiry, and CS concepts. The following high-leverage teaching practices enable students to *equitably* participate in student-led *inquiry* around important *CS concepts*: (a) provide a meaningful context for learning; (b) scaffold the development of CS concepts; (c) facilitate peer inquiry and collaboration; and (d) encourage multiple forms of expression [12, pp 7-8]. Inclusiveness is supported by focusing on ideas that are meaningful to students, and activities in the curriculum provide space for teachers to incorporate students' background and culture. In addition, many activities focus on real-life issues in the community—for example, students can make games that communicate messages about healthy eating or about the plight of undocumented students [18]. Resting on equity are inquiry-based activities in which students are "expected and encouraged to help define the initial conditions of problems, utilize their prior knowledge, work collaboratively, make claims using their own words, and develop multiple representations of particular solutions." [18] By engaging students in equitable inquiry through the first two strands, students gain access to the domain content of computer science, the third strand.

3 TRANSLATING THE ECS CURRICULUM INTO CLASSROOM TEACHING

Curriculum materials and activities represent one component of the ECS program. Given the significant shift in the nature of computer science teaching required for successful implementation of ECS, teachers need extended professional development to successfully adapt to the ECS model of teaching [13]. The ECS professional development program is intentionally designed to prepare teachers to implement the inquiry-based activities while also guiding them to build a classroom culture that's inclusive of all students [13]. Professional development begins with a weeklong summer workshop prior to implementing ECS. There are five key components of the ECS professional development model, the first being that teachers engage in the process of collaborative inquiry in small groups in the same way that students will engage in inquiry. The second component is that, throughout the first week, teachers participate in inquiry specifically through a teacher-learner-observer model. Each small group is assigned a lesson in which the group co-plans and teaches the lesson to the rest of the participants, who experience the lesson as learners. After the lesson, all the participants engage in reflective discussion about the experience from the point of view of the three ECS teaching strands (equity, inquiry, and CS content). These first two components of ECS professional development are consistent with what Desimone and Garet [5] call active learning in professional development. Their review of professional development found that active learning was an important component of professional development as it significantly influenced changes in teacher practices.

The third component of ECS professional development is explicit discussion and reflection on equitable practices. During the workshop, the teachers read sections of *Stuck in the Shallow End* [16], which provides rich case study descriptions of the roots of inequity in computer science. The fourth and fifth components of ECS professional development are meant to sustain teacher development over long time spans, which is another key dimension of effective professional development [5]. The fourth component is ongoing professional development during the school year and a second weeklong workshop the summer after their first year of implementation. The fifth component of ECS professional development is the development of a professional learning community. It begins in the summer workshop through the formation of small groups that engage in collaborative inquiry. It's also built up through the trust that teachers develop as they engage in tough, open discussions about equity as well as through open, honest feedback on lesson design and implementation during the workshops.

In prior research, we examined whether teachers were able to translate what they learned in professional development to create meaningful experiences for students. We used a researcher-developed end of course survey, which focused on students' perceptions of the relevance of the course (α=0.67). Almost three-fourths of the students (71%) rated the course as highly relevant. These student ratings of the perceived relevance predicted the students' attitudes towards computer science. In this study, we seek to expand on the initial evidence using instruments that have established convergent validity with measures of teaching practice. The Tripod 7C [11] is a survey of student course experiences that was empirically validated with the Danielson Framework for Teaching as part of the Gates-funded Measurement of Effective Teaching project. Teachers from a variety of subject areas and grade levels were observed and scored using the Danielson Framework. The students of those teachers were surveyed about their course experiences with the teacher using the Tripod 7C survey. The dimensions of the Tripod were correlated with the scores from the dimensions on the Danielson Framework to validate the ways in which students experience the various dimensions of the Danielson Framework. There are seven dimensions of the Tripod 7C: Challenge, Control, Care, Confer, Captivate, Clarify, and Consolidate. These dimensions are well aligned to the three ECS teaching strands.

4 INSPIRING STUDENTS IN COMPUTER SCIENCE

For this research, we seek to build on our prior work by using the expectancy-value-cost model [1] as a mediator for predicting student learning. The expectancy-value-cost model is an extension of the expectancy-value model, which is based on decades of research conducted by Eccles [8,9] on students' choices of majors and careers. These choices are dependent on how much value students put in the field as well as their expectation that they'll be successful. Eccles' research has shown that over time, students' expectations for success are based on successful experiences with relevant school subjects. The value that students place on a particular field is influenced by their enjoyment of experiences in the field, perceptions of whether the field will meet personal goals, and the extent that the field is valued by family, friends, and educators.

Of the corpus of research on the link between expectancy-value and future aspirations, there is one study in particular that is directly related to this research. The study investigates how pedagogical approaches support growth in expectancy-value [23]. The study took place at three middle schools in Greece where students were just finishing their first year of instruction in information technology. The students were surveyed on their expectancy-value as it relates to information technology, as well as the extent to which their teachers used practices that made meaningful connections to the real world through active learning and student collaboration. These practices are similar to the equity and inquiry strands of ECS. The results indicate that exposure to meaningful experiences significantly predicted growth in the value dimension but not the expectancy dimension, providing support for the hypothesis that experiences in ECS could increase the value students place on computer science by engaging them in meaningful tasks.

5 ASSESSMENT OF COMPUTATIONAL THINKING PRACTICES

To measure the development of computational thinking practices, we used assessments that were aligned to the computational thinking practices in ECS [22]. The assessments were developed and field tested by SRI International over two years using Evidence-Centered Design (ECD) [20], an assessment methodology that is especially advantageous when the knowledge and skills to be measured involve complex, multistep performances. The ECD process involved (1) working with various stakeholders to identify the important computer science skills to measure, (2) mapping those skills to a model of evidence that can support inferences about those skills, and (3) developing tasks that elicit that evidence [2]. The assessments were field tested with 941 students over two years [21].

Separate pretest and posttest forms were created. The pretest contains six tasks that measures students' initial understanding of CS concepts and computational thinking practices. For example, in one task students are asked to develop an algorithm to assign students to after school clubs that maximizes students' preferred choice while keeping the enrollment within the limits for each club. There are a series of subtasks that ask students about specific aspects of the problem and the solution. Across the six tasks there are a total of 19 subtasks that are scored independently. The posttest also contained six tasks, two of which were on the pretest and four of which were different. The two common tasks were used to equate the two forms and allow for measurement of growth from pretest to posttest. SRI developed scoring rubrics with student work examples for each of the tasks. Across all of the pretest and posttest tasks, there are a total of 30 question prompts that are each scored individually.

There were three sources of evidence for establishing the validity of the assessments at measuring the computational thinking practices covered in the ECS curriculum [3]. (a) Content validity was established through an expert review of the alignment between the knowledge and skills, the curriculum learning goals, and computational thinking practices. (b) Cognitive think-aloud interviews were conducted with a subset of students participating in the pilot test of the assessments. (c) The reliability of the assessments was moderate to high. The tasks within each assessment are well aligned with each other and with the targeted learning goals. See Snow, Rutstein, Bienkowski, and Xu [21] for additional details on the pilot study and validation results.

6 METHODS

Teachers who had previously participated in the ECS professional development program and were implementing ECS during the 2016–17 school year were invited to participate in the research. There were twenty ECS teachers from Chicago and twenty-seven ECS teachers from Wisconsin who accepted the invitation to participate. We were able to collect pre and post assessments and surveys from the students of eight of those teachers in Chicago and twelve in Wisconsin. The remaining teachers were dropped from the study since they provided only partial data.

There were 906 students who completed the ECS course, agreed to be in the study, completed the surveys and assessments, and whose parents consented for their participation. Table 1 shows the demographic characteristics of these research participants relative to the student demographic characteristics for ECS in CPS and total student population in Wisconsin. Students could select more than one race/ethnicity, so the percentages add to more than 100%. In both Wisconsin

Table 1: Demographic characteristics of ECS research participants

Demographics	CPS Research	CPS ECS	ECS WI	WI
Female	40%	50%	22%	50%
Caucasian	43%	12%	75%	71%
African-American	6%	29%	8%	9%
Hispanic	49%	50%	9%	11%
Asian	10%	7%	7%	4%

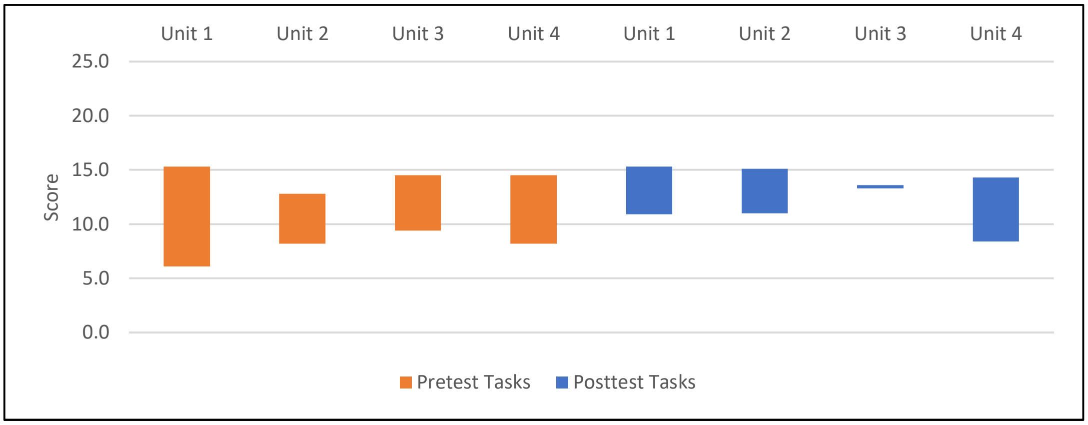

Figure 1: Range of item difficulties for pretest and posttest performance tasks

and CPS, there were fewer females than males who were taking an ECS course with the participating research teachers. With the graduation requirement in CPS, thepercentage of female participants is twice as high as for the teachers in Wisconsin.

The racial/ethnic characteristics of the participating students in Wisconsin were reflective of the demographic characteristics of students in the state. In Chicago, the demographic characteristics of the research participants were somewhat skewed relative to the population of students who completed ECS. There were more Caucasian students and fewer African-American students. Given the large sample size, there were sufficient numbers of students in each demographic category to be able to investigate differences in outcomes based on gender and race/ethnicity.

6.1 Assessments

During the 2016–17 school year, teachers administered the SRI-developed ECS pretest at the beginning of the year and the posttest at the end of the year. We hired The Graide Network for scoring the pretests and posttests. The Graide Network recruited and trained 26 undergraduate preservice teachers to score the performances tasks. They were provided training on each of the rubrics prior to scoring. As part of the training, each scorer scored a common set of 80 pretest responses from each question prompt in order to equate the severity of the scorers. For the posttest, we had overlapping subsets of scorers rate the same students. We used the *Facets* software version 3.71.4 to conduct Many-Facet Rasch Measurement analysis (MFRM) [10] to scale the student responses at each administration. *Facets* develops a model based on how well the student performed across the range of question prompts with set difficulties taking into account the severity of the scorer relative to the other scorers. Within MFRM, the goal is not for scorers to arrive at agreement on the scores, but instead to model the variation in how the scorers interpreted the rubrics. As long as the raters are internally consistent in how they apply the rubric, *Facets* can adjust the students' scores based on the severity or leniency of the scores relative to other scorers. We used the pretest tasks as the benchmark for scaling item difficulty. For scaling of the posttest scores, the item difficulties of the two common tasks were fixed based on the pretest scales. The overall model fit of the students, tasks, and scorers at each administration was high. For ease of interpretation, the logit scale produced by *Facets* was converted to a scale ranging from 0 to 25. Figure 1 shows a graphical representation of the range of difficulties for question prompts related to each unit. The bars show the range of difficulties for all of the subtasks related to each of the units. The overall difficulty ranges were similar across units.

6.2 Expectancy-Value-Cost

In this project, we used a validated, shortened version of Expectancy-Value-Cost survey. Barron and Hulleman [1] created separate middle school and high school versions of the survey questions. They conducted an extensive factor analysis to pare the survey length down to the shortest possible length that still provides high levels of reliability and construct invariance. In doing so, they discovered that the best factor structure treats Cost as a separate construct rather than as a negative valence within the Value dimension. The resulting survey instrument takes less than 10 minutes to administer so it can be administered more frequently. The Expectancy-Value-Cost scales were administered on the student survey at the beginning and end of the year. The alpha reliability of the Expectancy, Value, and Cost scales were 0.88, 0.92, and 0.83 respectively. We used the *Facets* software version 3.71.4 to combine the three scales into one EVC index score. The cost scale was reverse coded. We used the combined EVC index score to predict student learning outcomes.

6.3 Teaching Quality

As a proxy for the alignment of teaching practices to the ECS teaching strands, we used a combination of the Tripod 7C [11] and the pedagogical survey used by Vekiri [23]. The combination of scales across these two surveys provide a good approximation of the teaching practices that foster equity, inquiry, and development of computer science concepts. On the Tripod 7C, there are 36 Likert items that correspond to seven dimensions of teaching: *Care* relates to whether the teacher develops supportive relationships with students and is attentive to their feelings. *Challenge* refers to the extent to which the students acknowledge that the teacher places high expectations for rigor and performance. *Captivate* refers to the extent to which the teacher stimulates interest in the lessons. *Control* is related to the degree to which the class is both well-behaved and that the teacher is able to manage the class in such a way that learning can occur. *Confer* refers to the extent to which the

Table 2: HLM Model results for the student posttest scores by student and teacher characteristics

Posttest	Coefficient	Standard Error	t-test	p value
Average	**15.07**	**0.55**	**t(16) = 27.22**	**p<0.001**
Student Characteristics				
Pretest	**0.33**	**0.04**	**t(51) = 7.98**	**p<0.001**
Female	0.22	0.14	t(826) = 1.52	p=0.128
Hispanic	-0.03	0.19	t(479) = -0.18	p=0.857
Asian	0.09	0.28	t(826) = 0.34	p=0.734
African-American	0.001	0.31	t(826) = 0.005	p=0.996
Caucasian	0.07	0.18	t(477) = 0.39	p=0.694
Post EVC	**0.36**	**0.09**	**t(195) = 4.00**	**p<0.001**
Teacher Characteristics				
ECS Teaching Experience	**0.56**	**0.26**	**t(16) = 2.18**	**p=0.045**
Teaching Quality Index	-0.33	3.63	t(16) = -0.09	p=0.928

teacher elicits ideas from students and supports student discussion. *Clarify* refers to the extent that students feel that the teacher explains concepts well. *Consolidate* refers to the extent to which the teacher makes the learning experiences coherent for the students, giving feedback, and checking for understanding.

On the Vekiri pedagogy survey, there were 12 Likert items that correspond to three dimensions: *Active learning* relates to whether the teacher encourages exploratory and active learning through design and inquiry-oriented activities. *Collaboration* relates to whether the teacher created opportunities for student interaction and collaboration. *Meaningful learning* relates to whether the teacher highlighted the applications of computer science to the real world and tried to make computer science relevant to students' interests and everyday life.

The students within each class completed overlapping subsets of these scales on the end of course survey. We used *Facets* to develop a scale score for each student and then combined students' scores into a Teaching Quality Index score for each teacher using *WHLM* software version 7.24q.

7 RESULTS

In order to investigate the extent to which students' course experience and their attitudes towards computer science influence the development of computational thinking practices, we developed a series of models to first test the growth of computational thinking practices from pretest to posttest and then test which factors influence the amount of growth, including demographic characteristics, student attitudes, and teaching practices. The average pretest score was 15.2 out of 25 and the average posttest score was 17.0 for a growth of almost two points. We used a paired t-test to determine that this growth was statistically significant ($t(905)=20.3$, $p<0.001$) with a large effect size of 0.74, adjusted for the correlation between the pretest and posttest.

In the next model, we investigated the extent to which students' course experience and their attitudes towards computer science influence the development of computational thinking practices. We used the index of teaching quality and the years of experience teaching ECS as a proxy for students' course experience. In addition, we examined whether there were differences in student performance by students of different gender and racial/ethnic backgrounds. Table 2 shows the results of the analysis. Since students are nested within teachers, we conducted a Hierarchical Linear Model using *WHLM*. Students' demographic characteristics and the Expectancy-Value-Cost (EVC) index were included at student level. At the teacher level, we included the number of years the teacher had taught ECS and the index of teaching quality. After controlling for pretest performance, there were no statistically significant differences in posttest performance by gender or race/ethnicity. There were two statistically significant factors: the end of course EVC index and the number of years of experience teaching ECS. The teaching quality index did not have a direct effect on student posttest performance.

We examined the extent to which student demographic factors, the teaching quality index, and years of ECS teaching

Table 3: HLM Model results for the student post EVC scores by student and teacher characteristics

Post EVC	Coefficient	Standard Error	t-test	p value
Average	**2.68**	**0.11**	**t(16) = 23.46**	**p<0.001**
Student Characteristics				
Pre EVC	**0.65**	**0.04**	**t(72) = 15.57**	**p<0.001**
Female	-0.02	0.06	t(28) = -0.39	p=0.703
Hispanic	**-0.18**	**0.07**	**t(242) = -2.67**	**p=0.008**
Asian	0.01	0.11	t(118) = 0.09	p=0.932
African-American	-0.21	0.12	t(98) = -1.77	p=0.080
Caucasian	0.09	0.07	t(182) = 1.35	p=0178
Teacher Characteristics				
ECS Teaching Experience	0.02	0.04	t(16) = 0.46	p=0.649
Teaching Quality Index	**2.53**	**0.61**	**t(16) = 4.14**	**P<0.001**

experience influenced post expectancy. We once again conducted a Hierarchical Linear Model using *WHLM*. Table 3 shows the results of the analysis. There was one statistically significant difference in post EVC by gender or race/ethnicity. Hispanic students had lower levels of EVC after controlling for pre levels. Years of ECS teaching experience did not have a statistically significant effect on post EVC. The teaching quality index did have a statistically significant direct effect on students' levels of post EVC, which in turn influenced student learning outcomes.

8 CONCLUSION

A primary goal of the *Exploring Computer Science* curriculum and professional development program is to contribute to broadened participation of women and minorities and increased equity in the field of computer science. The curriculum is composed of activities that are designed to engage students in computer science inquiry around meaningful problems in the context of a classroom culture that's culturally relevant and inclusive of all students. In this research we examined the outcomes of implementations of ECS in Wisconsin and Chicago. If teachers are successful at implementing ECS in a manner consistent with the curriculum model, we hypothesized that the course experience would have a positive effect on students' expectancy-value and improve student learning outcomes. Overall, students achieved large learning gains from pretest to posttest and those learning gains were spread equitably across gender and race/ethnicity. These learning gains were larger for teachers with more experience teaching ECS, which suggests the extended professional development program provides opportunities for teachers to continue to improve their practices. These results also imply the importance for principals to maintain consistency in ECS teaching assignments from one year to the next.

In this research, we incorporated validated measures of teaching practices that are aligned to the equity, inquiry, and CS concepts teaching strands in ECS. The extent to which students experienced these practices positively influenced their expectancy-value toward computer science, which in turn had a positive impact on student learning. However, after controlling for these teaching practices, Hispanic students had lower levels of expectancy-value. Future research will examine how teachers' implementations of ECS can be better customized to meet the needs of Hispanic students. Despite this ethnic difference in attitude, these results provide evidence that the teachers' emphasis on equity within ECS is contributing to the broadening of participation in computer science. Eccles' prior research on expectancy-value suggests that the overall increased expectancy-value derived from the ECS course experience will lead to increases in the number of students who pursue computer science.

ACKNOWLEDGMENTS

This work was supported by grants CNS-1339392, DRL-1418149, CNS-1543217, CNS-1738515, CNS-1738572, CNS-1738691, and CNS-1738776 from the National Science Foundation (NSF). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

REFERENCES

[1] Kenneth E. Barron and Chris Hulleman. 2015. Expectancy-Value-Cost Model of Motivation. In International Encyclopedia of the Social & Behavioral Sciences (2nd ed.), James Wright (Ed.). Elsevier, 503–509.

[2] Marie Bienkowski, Eric Snow, Daisy Rutstein, and Shuchi Grover. 2015. Assessment Design Patterns for Computational Thinking Practices in Secondary Computer Science: A First Look. Technical Report. SRI International, Menlo Park, CA.

[3] Robert L. Brennan (Ed.). 2006. Educational Measurement. Rowman & Littlefield.

[4] Rhonda Christensen, Gerald Knezek, and Theresa Overall. 2005. Transition Points for the Gender Gap in Computer Enjoyment. Journal of Research on Technology in Education 38, 1 (2005), 23–37.

[5] Laura M. Desimone and Michael S. Garet. 2015. Best Practices in Teachers' Professional Development. Psychology, Society, and Education 7, 3 (2015), 252–263.

[6] Lucia Dettori, Ronald I. Greenberg, Steven McGee, and Dale Reed. 2016. The Impact of The Exploring Computer Science Instructional Model in Chicago Public Schools. *Computing in Science & Engineering, 18,* 2 (2016), 10-17.

[7] Oliver Dickhäuser and Joachim Stiensmeier-Pelster. 2003. Gender Differences in the Choice of Computer Courses: Applying an Expectancy-Value Model. Social Psychology of Education 6, 3 (2003), 173–189.

[8] Jacquelynne Eccles. 2009. Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. Educational Psychologist 44, 2 (2009), 78–89.

[9] Jacquelynne S. Eccles. 2005. Studying Gender and Ethnic Differences in Participation in Math, Physical Science, and Information Technology. New Directions for Child and Adolescent Development 110 (Winter 2005), 7–14.

[10] Thomas Eckes. 2011. Introduction to Many-Facet Rasch Measurement. Peter Lang.

[11] Ronald F. Ferguson and Charlotte Danielson. 2014. How Framework for Teaching and Tripod 7Cs Evidence Distinguish Key Components of Effective Teaching. In Designing Teacher Evaluation Systems, Thomas J. Kane, Kerri A. Kerr, and Robert C. Pianta (Eds.). John Wiley, 98–143.

[12] Joanna Goode and Gail Chapman. 2016. Exploring Computer Science (Version 7.0). http://www.exploringcs.org/curriculum (2016).

[13] Joanna Goode, Jane Margolis, and Gail Chapman. 2014. Curriculum is Not Enough: The Educational Theory and Research Foundation of the Exploring Computer Science Professional Development Model. In SIGCSE '14. Association for Computing Machinery, 493–498.

[14] Pam Grossman, Karen Hammerness, and Morva McDonald. 2009. Redefining teaching, re-imagining teacher education. Teachers and Teaching: Theory and Practice 15, 2 (2009), 273–289.

[15] Mizuko Ito, Heather Horst, Matteo Bittanti, danah boyd, Becky Herr-Stephenson, Patricia G. Lange, C. J. Pascoe, and Laura Robinson. 2008. Living and Learning with New Media: Summary of Findings from the Digital Youth Project. John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning. (Nov. 2008). http://www.macfound.org/media/article_pdfs/DML_ETHNOG_WHITEPAPER.PDF

[16] Jane Margolis, Rachel Estrella, Joanna Goode, Jennifer Jellison Holme, and Kimberly Nao. 2010. Stuck in the Shallow End: Education, Race, and Computing. MIT Press.

[17] Jane Margolis and Allan Fisher. 2002. Unlocking the Clubhouse: Women in Computing. MIT Press.

[18] Jane Margolis, Jean J. Ryoo, Cueponcaxochitl D. M. Sandoval, Clifford Lee, Joanna Goode, and Gail Chapman. 2012. Beyond Access: Broadening Participation in High School Computer Science. ACM Inroads 3, 4 (Dec. 2012), 72–78. https://doi.org/10.1145/2381083.2381102

[19] Steven McGee, Randi McGee-Tekula, Jennifer Duck, Ronald I. Greenberg, Lucia Dettori, Dale F. Reed, Brenda Wilkerson, Don Yanek, Andrew M. Rasmussen, and Gail Chapman, G. 2017. Does a Taste of Computing Increase Computer Science Enrollment? *Computing in Science & Engineering, 19*, 3 (2017), 8-18.

[20] Robert J. Mislevy and Geneva D. Haertel. 2006. Implications of Evidence-Centered Design for Educational Testing. Educational Measurement: Issues and Practice 25, 4 (2006), 6–20.

[21] Eric Snow, Daisy Rutstein, Marie Bienkowski, and Yuning Xu. 2017. Principled Assessment of Student Learning in High School Computer Science. In ICER '17 Proceedings of the 2017 ACM Conference on International Computing Education Research. 209–216.

[22] Eric Snow, Carol Tate, Daisy Rutstein, and Marie Bienkowski. 2017. Assessment Design Patterns for Computational Thinking Practices in Exploring Computer Science. Technical Report. SRI International.

[23] Ioanna Vekiri. 2013. Information science instruction and changes in girls' and boy's expectancy and value beliefs: In search of gender-equitable pedagogical practices. Computers & Education 64 (May 2013), 104–115.