A Mid-Project Report on a Statewide Professional Development Model for CS Principles

Jeff Gray
University of Alabama
Department of Computer Science
Tuscaloosa, AL
gray@cs.ua.edu

Kathy Haynie
Haynie Research and Evaluation
Skillman, NJ
kchaynie@stanfordalumni.org

Sheryl Packman
Haynie Research and Evaluation
Skillman, NJ
sheryl.packman@gmail.com

Mary Boehm
A+ College Ready
Birmingham, AL
marydboehm@bellsouth.net

Carol Crawford
A+ College Ready
Birmingham, AL
carolk.crawford@gmail.com

Deepa Muralidhar
Atlanta Girls School
Atlanta, GA
deepa.muralidhar@gmail.com

ABSTRACT

This paper summarizes our mid-project results in developing and evaluating a CS Principles (CSP) professional development (PD) model for training several cohorts of teachers across an entire state geography. CS4Alabama is an NSF-funded project that has adopted the successful practices of a national AP training program developed by the National Math and Science Initiative (NMSI) in pursuit of scalable deployment and sustainable persistence of new CSP courses across Alabama. We have created a CSP curriculum and PD program based on year-long in-person training and distance learning collaboration, which also was offered as a national MOOC as part of Google CS4HS. A statewide Teacher Leader (TL) model is used, where those who have previously taught rigorous CS courses serve as mentors in training new peer cohorts as they establish CSP courses in their schools. Teachers in these cohorts collaborate together on content and pedagogical learning experiences, fostered by the TLs. This paper reports on assessment results that are uncovering the facets of our model that are most suitable for building a sustainable network of CSP teachers. The paper summarizes our PD model, offers various lessons learned, and details the findings of the project's external evaluation team.

Categories and Subject Descriptors

K.3.2 [**Computer and Information Science Education**]: Computer science education

Keywords

CS Principles; Professional Development

SIGCSE '15, March 4–7, 2015, Kansas City, MO, USA
Copyright 2015 ACM 978-1-4503-2966-8/15/03…$15.00
http://dx.doi.org/10.1145/2676723.2677306

1. INTRODUCTION

As the interest in the new CS Principles (CSP) course grows as it approaches the 2016-2017 inaugural offering by the College Board, an obvious question has emerged: How will we train the mass of teachers needed to offer this course at scale? This paper describes our efforts on a new NSF CE21 proposal that addresses that question from the perspective of a state (Alabama) that has a traditionally low participation level on the current AP CS A exam. The ability to scale and sustain professional development in particular with teachers who have a STEM background but little CS content knowledge, will be vital to growing the CS10k vision.

The three-year participants of our project include: 1) a cohort of 50 high school teachers who will receive year-long professional development training to teach the CSP course; 2) students of those teachers, who will participate in summer camps, weekend study sessions, and statewide competitions; 3) the Alabama State Department of Education; 4) the principal investigators; 5) an evaluation team with extensive experience in evaluating CSP at a national level; 6) a K-12 teacher who has been a national Pilot participant and member of the CSP development committee; and 7) college students from secondary education and computer science majors who will assist with project tasks. The pedagogical approach is centered on inquiry/discovery-based techniques that introduce computer science as a broad set of topics, as defined by the learning objectives contained in the Big Ideas of CSP [5].

The goal of our project is to broaden participation in computing and especially to engage young women and other members of underrepresented minority (URM) groups. All of the efforts were carefully evaluated by an external team (Haynie Research and Evaluation), who provide a detailed view into a computing education intervention across the entire state of Alabama. Our Year 1 data sources include evaluations from over 300 students who completed a CSP course in 2013-2014, 9 secondary school teachers who attended PD sessions and taught the CSP course, and 23 secondary school teachers who began their preparations to teach the new course in 2014-2015. In this paper, we describe our PD model, results for participating instructors, results for implementations of the new CS course, and student course outcomes. We will describe how the first half of this project has broadened participation in Alabama. Several peripheral benefits have emerged as a result of this project, including:

- Alabama became the first state in the US to have CSP officially recognized as a statewide registered course
- a successful collaboration with our State Department of Education to have CSP and CS A count as one of the four math electives needed for high school graduation
- a heightened awareness of the importance of computer science at the highest levels of administration within our state's Department of Education
- the formation of the first CSTA chapter in Alabama

2. TEACHER PD MODEL

The CS4Alabama project aims to impact an entire state by preparing high school teachers to offer CSP courses through a replicable PD model that could be adopted in other states. We have openly made available all of our curriculum materials to anyone interested [1] and also developed online materials as part of a CS4HS MOOC offered throughout 2014 [2]. The specific details of our PD model include the following components:

- conduct year-round PD for a total of 140 hours per year, while training 50 teachers over three years in content knowledge and pedagogy related to CSP, including:
    - summer training (face-to-face and online)
    - bi-weekly meetings throughout the year
    - face-to-face weekend sessions each semester
- utilize the Master Teacher approach that was popularized by UTeach [3] and integrated into the NMSI model
- the Teacher Leader (TL) role allows the project to utilize the existing experience and skills among teachers who are already offering rigorous CS courses

The Logic Model for our project is in Figure 1, which shows a list of the inputs to the project, strategies for project implementation, and the expected short-term and long-term outcomes [4].

3. MID-PROJECT RESULTS

We are at the mid-point of our three-year project. This section describes pertinent evaluation results from the first full year of the project with our TLs, as well as the first half of our second year with a cohort of STEM teachers.

3.1 Year 1: Teacher Leader PD

In Year 1 of the project, ten TLs were recruited from our past contacts with those who we knew were already offering some form of CS at their school. This group of seasoned instructors has an average of 18 years teaching in grades K-12, and 12 years teaching CS at the high school level. All instructors have at least a master's degree, and most had taken at least 30 hours of PD in CS over the last three years. In the first year of the project, these teachers led the creation of teacher-driven lesson plans, pacing guides, and curriculum materials that formed the basis for mentoring of Cohort 2 teachers.

3.1.1 Summer 2013 PD

Summer training for the TLs took place during the last week of June 2013. The entire four-day period was rife with opportunities for developing and refining resources, networking, collaboration, and the building of friendships and team spirit among the teachers. It was evident throughout the PD workshop that a very special and powerful effort was underway to effectively scale up computer science in our state. Another major accomplishment that occurred during the meeting was the formation of a CSTA chapter and the election of officers.

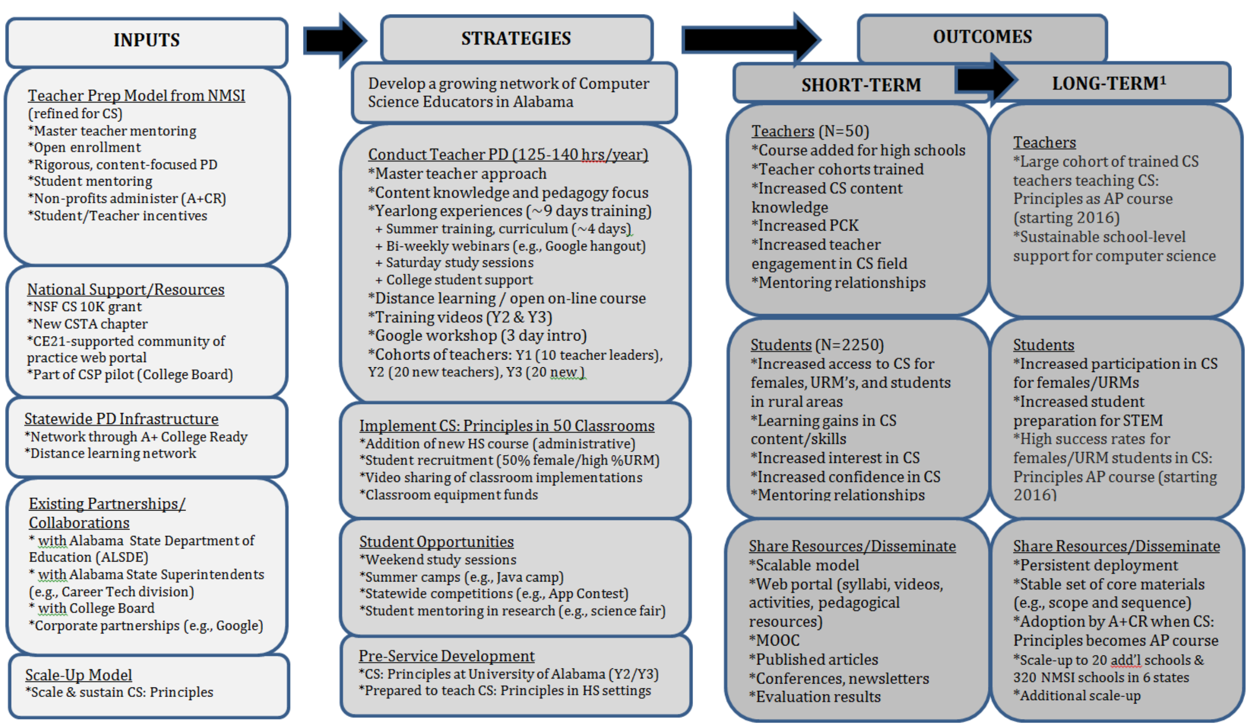

Figure 1: Logic Model for our CS Principles Project

The most critical aspects of the workshop were rated highly by the teacher participants: project overview, requirements of the CSP Performance Tasks, support for syllabus and lesson plan development, discussion of lesson plans, and opportunities for collaboration. High ratings were given to the facilitators' knowledge and effectiveness, the quality and planning of the workshop, the appropriateness of the activities, the enthusiastic and engaging atmosphere, and gaining knowledge and skills/strategies useful for instruction. At the end of the training, instructors wanted more resources and ideas to teach the material, further opportunities for collaboration, and help with technical writing components of required Performance Tasks.

3.1.2 Fall Weekend Session

An all-day PD session took place on October 19, 2013, attended by eight of the TLs. Overall, the session covered several project topics, including pacing guides, common assessment, and topics related to the future of the project. The most focus was given to discussion of the Data Performance Task and the Data Learning Objectives. Overall participant ratings indicated that the workshop was a success in terms of: quality of workshop planning, engagement, pacing, instruction, workshop materials, collaboration, teaching preparation, and learning about Data. Following the one-day session, participants prioritized their next steps as: studying the resources, planning for another section of the course, modifying lesson plans for Big Data, revising the fall pacing guide and implementing what they learned in their classes.

3.1.3 SIGCSE 2014

From March 6-8, 2014, most of the TLs attended the SIGCSE conference, funded by the project. In addition to the poster session presented by our project, the majority of the teachers attended 4-6 sessions, selected on the basis of relevancy and applicability of teaching the CSP course. Our TLs also presented a poster that summarized their first-year experiences, which provided them with a positive experience engaging with questions and conversations, including sharing successes, and helping others who are planning to teach CSP in the future (e.g., teacher recruitment and training). From a post-conference survey, the TLs found their experience at SIGCSE "very good" at developing understandings in content knowledge, tools and applications, curricular materials, and the larger context of CS Education. In addition, teachers mentioned learning about cooperative learning strategies, how to make our CSTA chapter stronger, and engaging girls in computer science. Their SIGCSE experiences helped them gather ideas for mentoring new teachers in Year 2 in terms of: lesson plan ideas and activities; choice of language(s) for CSP; student recruiting ideas; use of sample Performance Task rubrics; pacing guides; virtual trainings; and how to teach the writing aspect of Performance Tasks. The project also sponsored a special CSTA meeting between Alabama and Georgia teachers, providing an opportunity for Alabama teachers to form new collaborations.

3.1.4 Biweekly Meetings

Throughout the 2013-2014 academic year, project participants met every two weeks to discuss needed resources, lesson plan development, implementation successes and concerns (especially implementation of Performance Tasks), collaboration, project logistics, equipment, available tools, project evaluation, preparation for mentoring Cohort 2 teachers, and professional opportunities. These highly productive and amicable meetings focused on areas outlined in teachers' evaluations of PD sessions, as well as issues that arose in day-to-day course implementations. These meetings assisted in assessing the pulse of the project and continuing the collaborations across the TLs.

3.2 Year 2: Teacher Leader and Cohort 2 PD

In Year 2 of the project, twenty-three additional high school instructors were recruited. This group of instructors had an average of 9 years teaching in grades K-12, yet only four of the instructors had experience teaching computer science (1, 2, 2, and 3 years). A majority of instructors had earned masters degrees, with instructors citing degrees and/or certification in the following fields: mathematics education (43%), mathematics (35%), business (35%), engineering (26%), science or science education (22%), and computer science (13% - three of these four teachers all had over a decade of industry experience). Most instructors (74%) had taken 3 or fewer CS courses in college. Only two Cohort 2 teachers came from a school that already offered CS in 2013-2014. Figure 2 shows the geographic distribution of TLs from year one (orange/red) and the new Cohort 2 teachers (blue).

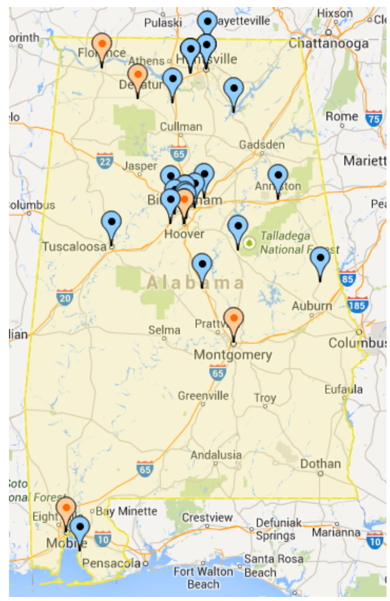

Figure 2: Locations of Teacher Leaders and Cohort 2 Teachers

3.2.1 Summer 2014 PD

Prior to the summer 2014 PD meeting, Cohort 2 teachers had an opportunity to learn the course content online. A MOOC was provided to these teachers, which was expanded to a national focus through support by Google CS4HS [2]. The online portion of the training provided content knowledge that would be built upon via the face-to-face teacher meeting in late June. The online content of the training summarized the CSP Curriculum Framework [5], offered lessons on important Learning Objectives of the CSP Big Ideas, facilitated virtual office hours through Google Hangouts, introduced the Performance Tasks, and engaged the participating teachers with CSP community leaders who offered virtual talks that are archived on YouTube.

The face-to-face summer training for the TLs and Cohort 2 teachers took place from June 23th- June 27th, 2014. A total of 9 TLs attended on June 23 and June 24; twenty-one Cohort 2 teachers attended the training from June 24-June 27. On June 24th,

Table 1: Self-Reported Cohort 2 Gains in Big Ideas and Computational Thinking Practices

	Average Before PD	Average After PD	Diff	ttest
Total for Big Ideas	*17.13*	*21.33*	*4.20*	*p=0.002*
1) Creativity	2.80	3.13	0.33	N.S.
2) Abstraction	2.10	2.87	0.73	p=0.002
3) Data and information	2.27	3.00	0.73	p=0.003
4) Algorithms	2.20	2.87	0.67	p=0.002
5) Programming	2.33	3.00	0.67	p<0.001
6) The Internet	2.67	3.27	0.60	p=0.017
7) Global Impact	2.73	3.20	0.47	N.S.
Total for CT Practices	*15.61*	*17.60*	*1.99*	*p=0.046*
1) Connecting computing	2.53	3.07	0.53	p=0.020
2) Creating computational artifacts	2.08	2.67	0.59	p=0.015
3) Abstracting	2.07	2.67	0.60	p=0.011
4) Analyzing problems & artifacts	2.47	2.93	0.47	p=0.034
5) Communicating	3.20	3.13	-0.07	N.S.
6) Collaborating	3.27	3.13	-0.13	N.S.

*1 = None, 2 = Novice, 3 = Apprentice, 4 = Expert

all teachers from both years were present. On this day, five working groups were formed that matched two TLs with five Cohort 2 teachers. The summer PD was rated very highly by all instructors. Not surprisingly, Cohort 2 instructors' overall ratings as to specific elements of the PD were slightly lower than those of the TLs; however, Cohort 2 instructors also gave themselves higher ratings in terms of gaining skills and knowledge as a result of the PD experience. Cohort 2 instructors rated themselves before and after the PD and the online study on the Big Ideas and Computational Thinking practices. Most of the gains found were significant, some highly significant (Table 1). The 4+point gain for the total Big Ideas score was highly significant, and gains for each of the Big Ideas were also significant, with exception of Creativity (Big Idea 1) and Global Impact (Big Idea 7). The almost 2-point gain for the total of Computational Thinking practices was significant ($p<.05$), as were the gains for four of the six individual Computational Thinking practices. The exceptions to this were communicating and collaborating. It is interesting that the significant gains were found in the most content-intensive areas (the areas we typically think of when considering computer science) and non-significant gains were found for broader or less content-intensive areas (i.e., creativity, global impact, communicating, and collaborating). These areas also had the highest ratings in their categories before PD, so there may have been a ceiling effect for some respondents.

Not surprisingly, after the summer 2014 PD, self-rated levels of preparation for TLs were much higher than those of Cohort 2 instructors. Combining both cohorts, major gains in perceived preparation were found particularly for cultivating students' interest in CS and implementation/classroom instruction. Highest gains (24 percentage points or more) were found for using Performance Tasks in CSP instruction, facilitating self-directed learning and project-based work, developing lesson plans, teaching CS to URM students, facilitating cooperative learning, and integrating principles of social justice into instruction. After the PD, most instructors felt unprepared to differentiate instruction for students with a variety of learning needs (e.g., English-language learners, students with learning disabilities).

Instructors felt that the best aspects of the PD were the flash talks that were offered by the TLs. These flash talks were 15 minute "lessons learned" or content/pedagogy suggestions offered by the TLs to the Cohort 2 teachers (these lessons are also available on our MOOC [2]). The teachers from Cohort 2 reported that the most helpful parts of the PD were the opportunities to work with the TLs, the community building, and the resources and recommendations for implementation. Cohort 2 instructors would improve PD to include more time with content and tools doing hands-on activities such as Snap! (and less time with pedagogy). Follow-up assistance was requested in terms of the regular online meetings, mentoring and direct support for Cohort 2 teachers, and more examples, activities, and lesson plans.

3.3 Attitudes and Self-Rated Understandings

The pre- and post-course student surveys included many identical items to gauge the impact of the course on key constructs related to success in computer science. Only students with matched pre/post data were included in these analyses. To equalize the impact of all TLs, a weighted sample was created with all courses having a sample of 20 students. The weighted sample was 68% male, 59% White, 72% non-URM, and 47% female or URM. In addition, 29% of students were in the 11^{th} grade, 28% in 12^{th} grade, 27% in 9^{th} grade, and 16% in 10^{th} grade.

3.3.1 Students' Attitudes

The pre- and post-course student surveys included 30 items to assess confidence, interest, and motivation for computer science as well as students' beliefs about computer science and its usefulness. All scales had acceptable reliability (0.78 – 0.93),

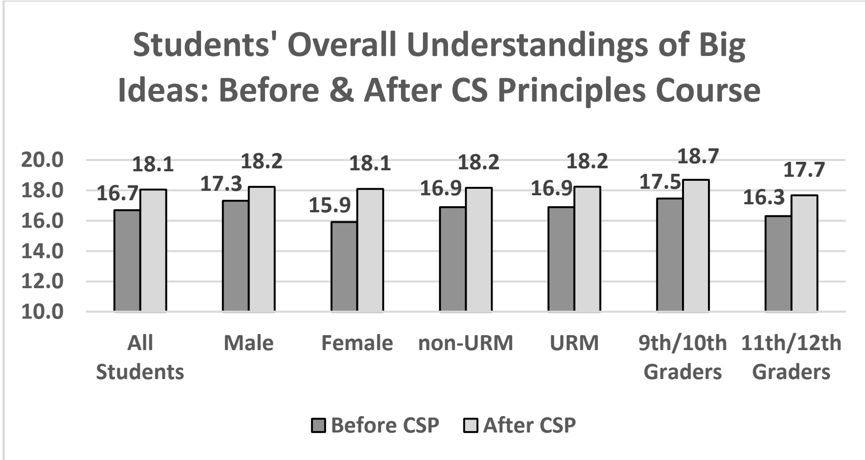

Figure 3: Analysis of Student Responses for Self-Efficacy

except for the beliefs scale. Results from the pre-course and post-course surveys were analyzed to determine if there were changes in confidence, interest, usefulness, motivation, and beliefs from the beginning to the end of the course. The only subscore that was significantly different from pre- to post-course was usefulness, which showed a decline. A repeated measures analysis of variance (ANOVA) was conducted to determine if there were significant mean differences by gender, URM status, and year in school. The only significant differences were found for interest and usefulness by year in school. In both cases, there were significant decreases for juniors and seniors with very small effect sizes.

3.3.2 Students' Understandings of Big Ideas

Students were also asked to rate their understanding of each Big Idea before and after taking the course. Responses were given on a 3-point Likert scale: 1=No Understanding, 2=Some Understanding, 3=Strong Understanding. The highest ratings were for Big Idea 7, Global Impacts of Computing, while the lowest ratings were for Big Idea 2, Abstraction. Scores went up for every Big Idea, with the greatest gains for Big Idea 2 (Abstraction) and Big Idea 4 (Algorithms).

A measure of students' overall understandings of the Big Ideas was the sum of the seven ratings, with a potential range of 0 to 21. A repeated measures ANOVA was conducted to determine if there were any pre/post differences in understandings for all students and for demographic subgroups. Overall understandings increased significantly for all students ($p<.01$). The only significant difference by subgroup was for gender whereby both males and females showed a significant gain from pre- to post-course, but that gain was greater for females (effect size=.03). In fact, females started the course with a lower overall knowledge of the Big Ideas and ended up almost identical to the males in their post-course ratings, which was an encouraging result.

3.3.3 Students' Self-Efficacy

Students rated their self-efficacy levels before and after the course in five areas directly related to the CSP Curriculum Framework [5]. Responses were entered on a 4-point Likert scale. The highest ratings were for "I am persistent at solving logic problems" and the lowest were for "I can write successful computer programs" Significant gains were found for "I can write successful computer programs," "I can effectively use abstractions and models to achieve goals," and "I can effective analyze the ethical, legal, and social implications of computing."

All self-efficacy items were summed together to create a total self-efficacy score for each respondent (with a potential range of 5 to 20). There was a significant increase in overall self-rated abilities from pre- to post-course ($p=.00$, effect size=0.12). A repeated measures ANOVA was conducted to determine if there were any differences by subgroup. The only significant finding was for URM status whereby non-URM students had significantly higher gains in self-efficacy ($p<.01$); however, the effect size was very small (0.04).

3.4 Common Assessment

A Common Assessment was developed by the TLs, who contributed items intended to assess six of the seven Big Ideas in the course: abstraction, algorithms, data, Internet, programming, and impact. Project leaders reviewed all items and selected 20 multiple-choice items - four items were selected for each subscale, except impact. Each instructor administered the common assessment online to their students within the last two weeks of their course.

To control for any teacher effects, each course implementation (N=9) was weighted to the same number of students (N=23) and only the first administration of the assessment was used for teachers with multiple course implementations. The total weighted sample of 207 students was 34% female, 25% URM, and 48% female or URM (a project goal is a sample of at least 50% female or URM students). A majority of students (57%) were in the 11th or 12th grade. The overall internal consistency reliability of the 20 items in the Common Assessment was 0.658; 19 of the 20 items exhibited acceptable item discrimination levels (>0.20).

For the total student sample, the mean total correct was 12.3 of 20 items (SD=3.3). There was a significant difference in performance by gender, whereby females (M=13.1) performed better than males (M=11.9) ($p<.02$; $\square^2=.03$). There was also a surprisingly significant difference in performance by year in school, whereby 9th and 10th graders (M=13.5) performed better than 11th and 12th graders (M=11.4) ($p=00$; $\square^2=.09$). There were no differences by racial/ethnic group or URM status.

4. DISCUSSION OF ASSESSMENT

This section discusses several of the most significant results of student learning from the first full year of the project and suggests a reflective approach to updating the project PD needs based on the formative assessment feedback [6].

4.1.1 Students' Self-Ratings

What is most promising in these findings is that the gender gap across all Big Ideas disappeared from pre to post. This is certainly a promising finding and suggests that the instructors of the classes with substantial proportions of female students (four classes were over 30% female) were successful in helping female students succeed, perhaps even relative to their male counterparts. Surprising, neither males nor females showed significant gains in confidence, interest, or motivation from pre to post. However, in addition to the learning gains for Big Ideas, students' self-efficacy levels increased in a number of areas including writing successful computer programs, and using abstraction and models. This suggests that students' beliefs in their abilities to do some key course-related activities were impacted more relative to their general perceptions of their confidence in computer science. Curiously, although there was not a significant gain in students' ratings of their knowledge of Big Idea 7, "Computing has Global Impacts," overall self-efficacy ratings significantly increased for the statement, "I can effectively analyze the ethical, legal, and social implications of computing." This suggests that the more specific wording may have elicited greater student recognition.

4.1.2 Student Performance on Common Assessment

While the common assessment results are certainly promising, some interpretation of these findings is needed. On the surface, it seems counter-intuitive that the 9th and 10th graders in our sample would significantly outperform the 11th and 12th graders. However, the two highest scoring classes of students were comprised almost entirely of 9th graders; these students also spent considerably more time on the common assessment, compared with the rest of the students. One of these high-scoring classes included one 10th grader and one 11th grader. The second class, from a technology magnet school, was entirely 9th graders. Please notice that the starting baseline (Figure 3) of understanding for these 9th/10th graders was also higher than the starting baseline for the 11th/12th grade participants. In addition, the class from the technology magnet school – by far the highest scoring school in our sample (averaging 1 to 5 more items correct than the other schools) – was 72% female. Further study indicated that this technology magnet school was ranked 8th in the state, out of hundreds of high schools. When this unique school was removed from our analysis, the average difference by grade level disappeared, and the "gender gap" was reduced by 23%.

With only nine participating schools in our Year 1 sample, it is difficult to generalize these student outcomes, since one school can greatly influence the aggregate results. However, considering this sample both without the outstanding magnet school, the females in the sample outperformed the males (the gender difference was largest for the Internet items), the under-represented minority students performed about as well as the non-URM students, and there was little or no difference between students in upper and lower grades.

5. CONCLUSION

We believe that the particular challenges that we face in our project (e.g., geographic displacement of teachers, wide range of teacher preparation and experiences, diversity in student and teacher populations) represent a microcosm of similar challenges that are emerging nationally to scale and sustain a new cohort of computer science teachers. The challenge of geographic displacement of teachers has driven the need for a blended model of professional development that has both online and face-to-face instruction in content preparation and pedagogical strategies. The diversity of teacher background and preparation has led us to the implementation of the TL model, whereby more experienced teachers help to create curriculum resources and lead the mentoring process. One of the aspects of our project that we are most excited about is that the majority of lesson plans associated with our project are all teacher-created, and address the specific needs of peer teachers and their daily challenges faced when introducing a new computer science course. Likewise, for the 2014-2015 school year, all of the online mentoring of biweekly Hangouts is led by pairs of TLs who are assigned groups of Cohort 2 teachers who have similar backgrounds (e.g., career tech teacher group, math instructor group). We believe that the empowerment and growth of the TLs will drive the sustainability of our efforts long after the funding for our project has ended.

There is still much work to do on our project. The mentoring and assessment of our model for Cohort 2 teachers continues throughout the project, and a third cohort of 20-25 teachers will be recruited for participation in initial PD for summer 2015. Although the challenges will continue regarding the lack of computer science content preparation among future teachers, we also believe that the TLs' experience and the growing richness of our curriculum resources will assist in expanding computer science in Alabama both in terms of quantity and quality.

REFERENCES

This project was funded in part by an NSF CE21 award (1240944) and several Google CS4HS awards.

REFERENCES

[1] CS4Alabama, http://cs4alabama.org/

[2] CSP4HS MOOC, https://csp-cs4hs.appspot.com/

[3] UTeach, http://changetheequation.org/featured-programs/uteach

[4] Michael S. Garet, Andrew C. Porter, Laura Desimone, Beatrice F. Birman and KwangSuk Yoon, "What Makes Professional Development Effective? Results from a National Sample of Teachers," American Educational Research Journal, Vol. 38, No. 4 (Winter, 2001), pp. 915-945.

[5] CS Principles Curriculum Framework, http://apcsprinciples.org/cs-principles-documents/

[6] Laura M. Desimone, "Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures," Educational Research, August 2009, 38. pp. 181-199.