CS10K COMMON DATA ELEMENTS
Evaluation Working Group Year 1 Project Report
August 2015

EXECUTIVE SUMMARY

The National Science Foundation's (NSF) CS10K program "aims to have rigorous, academic computing courses taught in 10,000 high schools by 10,000 well-prepared teachers. CS10K proposals will focus on high school computer science teachers, providing preservice and inservice teachers with courses, professional development opportunities, and long-term, ongoing support" (NSF Program Solicitation 15-537). This report is a first step towards answering three questions central to the National Science Foundation's CS10K program: 1) How many new teachers have participated in professional learning (PL) through CS10K-funded projects; 2) What are the demographic characteristics of those teachers; and 3) what teaching experience do these teachers possess?

Twenty of the 35 projects funded through CS10K grants reported data. Between the 2011-12 and 2014-15 academic years, 682 new teachers (678 new to CS10K teachers plus 5 that were "Returning" during the first year of the project) have participated in PL through the CS10K programs that provided data for this report. They teach in at least 463 public schools in 35 states and the District of Columbia. Slightly more female teachers entered CS10K projects than male teachers, and most identify as white and non-Hispanic. Thirteen of these reporting projects offer PL for Computer Science Principles (CSP) and seven offer PL for Exploring Computer Science (ECS). Most teachers enter professional learning with at least 6 years of K-12 teaching experience. Many projects provide professional learning to experienced, "master" CS teachers in the first year of the program and to less experienced teachers in subsequent years. Finally, this report also describes the data collection process and offers guidance for future data collection efforts.

An Evaluator Working Group (EWG) was convened at NSF's request to lead a community-wide common data collection effort. In collaboration with currently-funded CS10K projects, the team piloted a set of common measures for data collection across projects, and a spreadsheet for compiling all collected data, and process to assess the feasibility of a program-wide reporting system. This first-year report describes the number, demographic characteristics, and teaching experience of the teachers participating in professional learning. This report also documents what the EWG has learned during the first year about the data collected, how it is collected, and the context in which these projects operate.

The results of this study uncovered the complexities of the CS10K landscape and data collection within this community and within the context of school systems. This round of data revealed the following lessons many of which will be addressed in this upcoming year of data collection:

The data collected

1. This effort is the first time evaluators across the CS10K projects have been asked to collect and report consistent data across projects.
2. The EWG requested historical data, but those data may not have been collected and organized such that it could be integrated into this report.
3. This study defines a "new" CS teacher as one who is new to the CS10K project as opposed to a teacher returning to the project to receive additional professional learning. Historical data

Prepared by

Summer, 2015

rzarch@sagefoxgroup.com
tmklin@sagefoxgroup.com

SAGEFOX CONSULTING GROUP
AMHERST, MA • ATLANTA, GA

jxavier@sagefoxgroup.com

collection practices prohibit the EWG from counting only those teachers who are truly new to teaching computer science.

4. Capturing past CS teaching experience is difficult as teachers have different definitions and self-assessments of the past computing courses they have taught.
5. Collecting information on the students benefiting from the teacher's PL is challenging due to district Research Review Board (RRB)/Institutional Review Board (IRB) requirements and evaluation resource constraints.

Understanding the context

6. CS10K projects have different models for PL (single vs. multi-year) and different content foci (ECS and/or CSP and/or other). The first year of project funding is often, but not always, a planning year. The first cohort of teachers is sometimes a cohort of experienced, master teachers.
7. The CS10K program is one of several programs to offer computer science (CS) PL to teachers in the U.S. Others include Broadening Participation in Computing, CE21 projects (non CS10K), Math/Science Partnership Programs, DRK-12, Project Lead the Way, Code.org, and Industry efforts (such as Microsoft TEALS). These programs also vary in the approach and focus of the CS PL. To truly assess the number of teachers new to teaching CS requires counting teachers in non-CS10K training programs who are not also counted as being trained via CS10K programs (i.e., some CS10K projects, like Taste of Computing (ToC) in Chicago includes teachers who overlap with the ToC project and the Code.org run teacher trainings).

Table of Contents

INTRODUCTION AND MISSION

In June 2014, NSF Program Officer Jan Cuny hosted a meeting for CS10K evaluators in Arlington, VA, and one topic was the development of a CS10K common data collection effort. During this meeting, evaluators explored the tools and procedures for obtaining common participation and outcome data across the CS10K projects. As a result of this meeting, SageFox Consulting Group submitted a request for supplemental funds (NSF award #1228355) to convene an Evaluator Working Group (EWG), a small group of evaluators from the CS10K community, to collect and report common data elements.

The Evaluation Working Group

The participants in the Evaluator Working Group include:

Tom McKlin	SageFox Consulting Group (lead evaluator)
Jen Duck	The Learning Partnership
Kathy Haynie	Haynie Research and Evaluation
Kathleen Perez-Lopez	American Institute for Research (AIR)
Gary Silverstein	Westat
Sarah Wille	Outlier Research & Evaluation, CEMSE, U Chicago
Jeffrey Xavier	SageFox Consulting Group
Rebecca Zarch	SageFox Consulting Group

Alan Peterfreund and Daryl Chubin are serving as advisors for this project.

The group was charged with the following nine tasks:

1. Finalize a set of demographic items to be used on CS10K Professional Learning surveys for teachers.
2. Finalize a set of demographic items/categories for reporting student demographic data.
3. Finalize an agreed-upon set of background items (school name, courses taught, etc.) describing teacher participants in CS10K projects.
4. Provide technical assistance to other evaluators on incorporating demographic items into their data collection routines.
5. Report to NSF the number and demographics of new teachers in CS as a result of programmatic efforts.
6. Identify and share among the broader CS evaluation community other instruments, scales, constructs, and approaches that other evaluators might use.
7. Curate a set of research questions that the research and evaluation community might answer.
8. Form a set of recommendations for ongoing data collection among evaluators.
9. Meet annually with evaluators at NSF-funded meetings of CS10K PIs and Evaluators to share instruments, methods, and promising directions. This will also serve as an "induction" for evaluators new to CS and invite them to become active participants in the CS evaluation community.

This report will focus on the work completed as part of tasks 1-5. Tasks 6-9 will continue into the next project year.

Evaluation Working Group (EWG) Guiding Questions

The EWG posed seven critical questions that guide this report and future reporting of common data elements. Table 1 presents the questions along with potential data sources and the task to which each question relates.

Table 1: List of major questions to be answered through CS10K common reporting

Q#	Task #	Common Question	Data Sources
Q1	T7	How many high school CS teachers are there in the United States?	Potentially, CSTA; AP Central; Code.org.
Q2	T1, T3, T5	How many teachers have been trained in CS10K projects? Specifically: • How many teachers have been trained in CS10K projects since the program's inception? • How many teachers have been trained in the new CS10K projects?	CS10Kcommunity.org ; Annual reports submitted through RPPR and Fastlane; EWG Common Elements Data Shell: Teacher descriptives/demographics
Q3	T3, T5	How many CS10K teachers are teaching with the instructional materials/approaches/ curricula they have been trained to teach?	EWG Common Elements Data Shell[1]: contact information, Teacher descriptives/demographics
Q4	T7	How many CS10K teachers have been trained but are not teaching with the instructional materials/approaches/ curriculum they have been trained to teach?	EWG Common Elements Data Shell: Teacher descriptives/demographics
Q5	T3, T7	How many CS10K high schools are there?	EWG Common Elements Data Shell: School data from evaluators
Q6	T1	How diverse (race/ ethnicity/ gender/ disability status/ CS experience) are the teachers who have been trained through the CS10K program?	EWG Common Elements Data Shell: Teacher descriptives/demographics; NCES data
Q7	T2	How many and how diverse (race/ethnicity/gender/disability status) are the high school students reached through the CS10K program?	EWG Common Elements Data Shell: Student demographics/District student level data, Teacher's schools[2]; NCES data

[1] The common data shell is the tool used to capture data across projects and is described below under Development of Common Measures and Reporting Tool

We surmise these to be the most critical questions to inform CS10K efforts while acknowledging the variation among CS10K projects, most notably the type of CS preparation offered through PL. This report is a product of the effort to answer questions two to seven.

METHODS AND LIMITATIONS

The methods used by the EWG include a peer-facilitated process to define common descriptives and tracking process. As members of the CS10K community, we understand the nature of the CS10K projects, the feasibility of collecting specific items and the resources required for collecting this data.

CS10K programs fall under two major project types:

- Those that provide PL. These can either be for CSP, ECS or another rigorous approach to teaching CS.
- Those that don't provide PL. These are either research projects or resources for other projects.

Not knowing the details of each project, the EWG reached out to all funded projects. Table 2[2] shows the number of projects funded per year. Of the 35 funded projects, 14 focus on ECS and 21 focus on CSP. There are two of the 35 projects that are research-focused and don't provide PL.

Table 2: Projects awarded per year and number submitting common data by award year.

Award Year	Number of Awards	Projects reporting
2012	13	9
2013	10	5
2014	12	6

Upon embarking on this effort, the EWG was well-aware of the challenge associated with the timing of the CS10K awards, which are made in the fall and winter, and how to define cohorts and reporting periods. The EWG agreed that, for data collection purposes, projects should consider participants to be active on an annual basis, starting in June (roughly) with the summer professional development and teaching through the following academic year. Participants recruited during the spring are, therefore, not to be included in any common data collection until they have begun intensive PL during the summer months preceding the academic year in which they are to actually implement the CS program. This is especially relevant to projects funded in 2014, as most were structured to begin offering PL to participants, summer of 2015, and into the 2015 – 2016 school year. We present our analysis results in terms of these reporting periods that range from early summer of one year to the end of the following school year, with the exact dates being project-specific. For example, the reporting period 2013-2014 refers to data collected from projects about teachers who began their PL during summer 2013 and may have taught CS through the 2013-2014 school year.

This is the first time a common data collection effort was implemented with CS10K projects. We asked evaluators to report on teacher and student information for the most recently completed academic year and to provide data for previous school years, when possible. To orient evaluators to the goals of the

[2] This data includes project-submitted data through July 31, 2015.

EWG work and the categories of data requested, we provided several options for evaluators to learn more about the work, and what will be requested of them in future years. For example, the team developed and shared several "technical documents" with recommended survey items; we provided a 90 minute webinar in October of 2014; an informational meeting and face-to-face support at the 2015 CE21 PI meeting and; developed a list of frequently asked questions, posted on the CS10K Community site. The team also assigned a dedicated EWG liaison to work with each CS10K project. In mid-spring 2015 we sent evaluators individualized EWG Data Shells requesting draft data (i.e., in as close a form as possible to the parameters specified in the "technical documents") with the twofold intention of having evaluators look closely at the data collection request in preparation for their 2015 end-of-school-year survey administration (in case there were opportunities to add our recommended questionnaire items to their existing instruments) and to use any submitted data to refine our analysis plan. The timeline for this effort is found below in Table 3:

Table 3: EWG timeline

Activity	**Date**
CS10K Evaluation meeting hosted by Jan Cuny	June 2014
EWG meeting #1	September 2014
Data collection tool development	September – October 2014
Webinar	October 2014
Tool refinement	October 2014-January 2015
Presenting to community	January 2015
EWG data collection from CS10K sites	March- July 2015
Draft data submitted to EWG	May 2015
Final data submitted to EWG	July 2015
EWG meeting #2	July 2015

Although many of the projects funded in 2014 are providing professional learning to their first cohorts of teachers in Summer 2015 (and are thus were unable to provide firm numbers for this submission), we included the PIs and evaluators for these projects in this effort so that they could plan for the common data collection moving forward.

Development of Common Measures and Reporting Tool

At its first meeting in September, 2014, the EWG developed the "common data elements shell," a spreadsheet designed to help CS10K evaluators structure their data and deliver it to the EWG. The shell supports the collection of data on teachers, schools, and students. The most up-to-date data shell can be found on the CS10K community website. The data shell was designed for each project to provide aggregated data on the participating teachers, the students taught by these teachers and the schools in which teachers offered their new CS curricula to the EWG. It allows for projects to enter information by project year (school year) and accounts for teachers who are new to the CS10K project versus those returning to PL opportunities to participate in an additional year of training. The specific tabs of the data shell include:

1. Teacher descriptives/demographics (gender, race, ethnicity, disability status and use of training materials in teaching[3])
2. Teacher experience (background in K-12 and CS teaching)
3. Student-level data (student enrollment, student demographics, and student course pass rates)

[3] Use of training materials will be moved to teacher experience in the next iteration of data collection.

4. School data (information about schools in which participating teachers work)

The EWG created a technical document to accompany the common elements data shell. It contains survey items that the EWG requested CS10K evaluators use in their teacher surveys moving forward. Both the data shell and technical document were revised as a result of a feasibility survey sent to current CS10K evaluators, a meeting with the broader evaluation community through a webinar in October, 2014 (27 attendees), and at the January, 2015, CS10K meeting.

Demographic items asking teachers about race and disability status can be particularly sensitive, and thus challenging. In our recommended survey items that ask about race, we use the Government Accountability Office (GAO) recommended categories, and instructed evaluators to allow participants to "select all that apply" from the racial categories listed. We then asked evaluators to provide aggregate information about teachers, which can result in the total number of persons across categories being more than the total number of teachers participating. Collecting student-level data of this nature requires extensive work with district IRB/RRB offices and thus not feasible with the current funding allocations.

Further, collecting accurate disability data is difficult. Although we initially considered listing specific disabilities in a teacher survey item, and asking teachers to select any that applied, we ultimately decided to collect disability status as a "yes/no" response to protect privacy and minimize potential respondent discomfort. The EWG also hosted an information session with Richard Ladner and Lyla Crawford to better inform data collection specific to disability status of teachers and minors.

Each demographic item described above also contains an option in which respondents have the option of selecting "prefer not to answer," thus allowing a more complete picture of whether data simply was collected, or if project participants simply chose not to provide certain data.

Data Collection Process

Most EWG members are also evaluators of CS10K projects; therefore, the EWG anticipated many of the challenges projects would face in providing data and planned to mitigate these difficulties. To drill down further into these issues, in September 2014 we administered a survey to the CS10K evaluators to ask about the feasibility of collecting different types of demographic and experiential information.

Based on this survey's results, the EWG prepared a webinar for project PIs and evaluators, held in October 2014. During the webinar we described the common data collection effort origins, plan, survey items, and data collection shell (see Figure 1). We shared an initial version of the data collection shell for evaluator feedback and advised participants that we would be requesting that draft data be

submitted six months later.

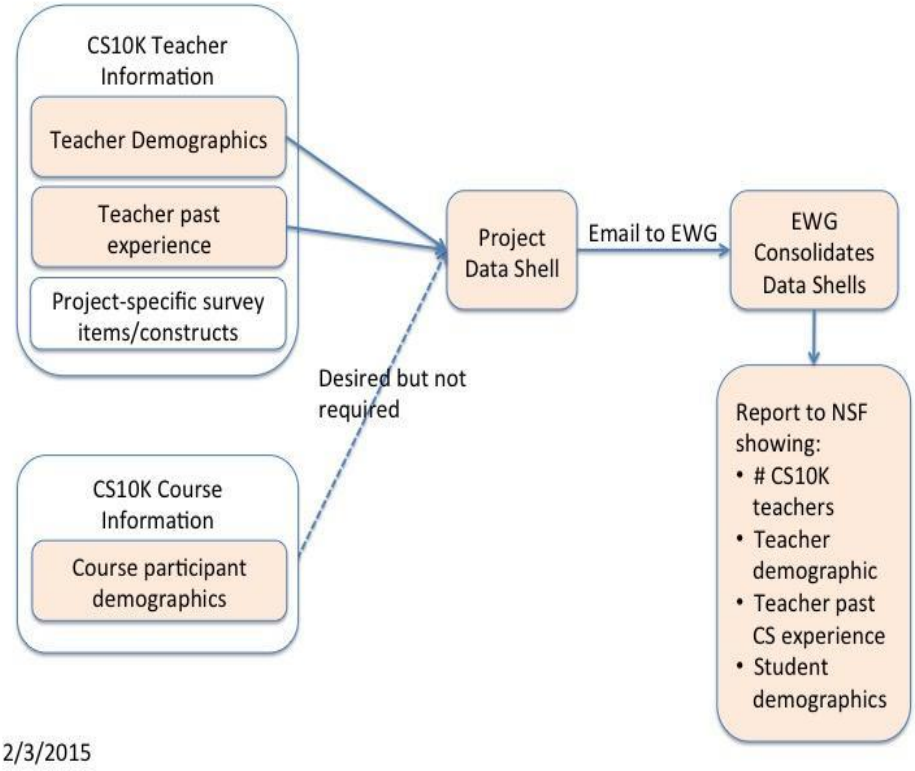

Figure 1. Process for aggregating CS10K project data

A link to the slides, webinar and notes from the call were posted on the CS10K Community website. Participants were also invited to the session planned for the 2015 CE21 PI meeting.

At the CE21 PI meeting, the EWG presented a detailed overview of the data collection shell and companion technical documents and responded to the community's questions. As anticipated, school district RRB/IRB concerns were at the forefront and participants shared ideas for working with district RRB/IRB offices to request or collect student-level data.

The EWG used a liaison approach to data collection. The eight members of the EWG, all of whom are active CS10K and/or Broadening Participation in Computing (BPC) evaluators or researchers, served as liaisons to evaluators in each CS10K project. We directly contacted each evaluator after the October, 2014 webinar to establish a connection, let the evaluators know that a data request was coming, and provided information and technical assistance on collecting and successfully completing the data shell for each project. Data collection formally started on March 13th, 2015 and ended on July 31st, 2015.

Participants / Respondents

With the exception of some of the most recently-funded projects, the EWG established contact with all CS10K projects to support the data collection effort. By the first week in August we received submissions from 20 of the 35 projects (see Table 2).

Not all projects submitted data and most that did submit provided only partial data:

- Some are research based and do not offer PL to teachers
- Some chose not to participate
- Some have yet to offer PL and thus do not have any data about teachers, students or schools. This is true for most of the 2014-funded projects, which offered their first PL in Summer 2015.

Most projects were able to report on the descriptives/demographics of the participating teachers and the schools in which the teachers implement the PL. Understanding teacher experience required adding EWG-recommended items to existing evaluator teacher surveys and collecting those data from teacher participants, and this may have presented a challenge for this pilot year of data collection (e.g., timing and instrument adjustments). The student-level (classroom) demographics information provided by evaluators was inconsistent (i.e., it was not clear how some data were collected), thus the EWG decided to exclude any student demographic information due to the district RRB/IRB requirements/constraints and uncertainty of how the data were collected.

The common data collection shell requires that evaluators provide data by project year for teachers *new* to the CS10K PL, and for those *returning* for a repeat year of PL. For analysis purposes, only counts for *new* CS10K teachers were included in the analysis to prevent duplicate teacher entries (and thus over-estimating the total number of teachers trained in ECS or CSP thus far). No teachers reported as "new" for the 2015-2016 year are included in our analyses, because they have yet to complete professional learning or any implementation of ECS or CSP, based on their PL experience. The 2015-2016 teachers will be included in next year's submission. The projects funded in 2014 were excluded because their teachers will not begin teaching ECS or CSP in the classroom until fall 2015. These projects were asked to participate in this pilot data collection year simply to familiarize them with the data collection effort and process to facilitate buy in and compliance in the 2015-16 reporting cycle.

Caveats

The EWG appreciates evaluators' responsiveness to this study. Based on the submissions, we know that the type of data requested varies in difficulty to collect. Historical data is particularly challenging for projects, and virtually impossible for items tied to survey items developed by the EWG during year one of the pilot. Though 20 projects submitted data, their reports vary greatly in completeness. See table 4 below for an overview of the number of projects submitting each type of data by year. Note that projects may have submitted partial data.

Table 4: Number of projects providing type of data by program year:

Period	Teacher Demographics	Teacher Experience	Classroom-level	School-level Data
2011-12	3	3	3	18
2012-13	5	5	4	
2013-14	11	11	11	
2014-15[4]	17	17	11	

[4] The projects funded in 2014 were asked to report on participants trained in 2015 to build a culture of participation, however they are not included in the analysis as these teachers have yet to implement the PL.

The common data elements shell asks evaluators to provide any caveats about the data so we could better understand the collection challenges and address them in future data collection efforts. Common challenges (by data shell tab) for each submission include:

1) Teacher demographics: Evaluators consistently reported teachers' gender, race and ethnicity. However, if the data were not previously collected, projects were unable to go back to teachers to collect demographic data, and if the race/ethnicity questions were asked differently from the recommendation provided by the EWG, findings should take this in to account.
2) Teacher experience: Projects were more likely to have data about past experience teaching in a K-12 environment and experience teaching CS in general, but were less likely to have documentation for the type of CS courses with which teachers were experienced.
3) CS Course Materials: Most projects train teachers on the use of ECS- or CSP- specific material, yet teachers' use of course-specific learning materials was less commonly reported than gender, race and ethnicity. One reason for the inconsistency is that the use of materials required a specific survey item that was likely not integrated into spring/year end teacher surveys for this pilot year. Further, some teachers are in a multi-year PL project (e.g. IDOcode) and may not have taught yet/will teach in the 2015-2016 year. The EWG plans to revise the existing survey item to better link professional learning to the course(s) for which the materials were created, so it is an even stronger item for the 2015-2016 school year.
4) Student participation: Collecting student-level data requires district RRB/IRB approval for administration of student questionnaire or collecting student data directly from the district. While we hoped to gather existing data on the number and demographics of students taught by program teachers, we came to understand that: a) projects collect these data in inconsistent ways (e.g. some gather ethnicity data while others do not) and b) the methods for providing the EWG with student demographic data may have been inconsistent, and in some cases may not have been acquired through required district RRB/IRB channels. That is, some student-level data in some sites may have been collected without explicit district review board approval and some demographic data (e.g., race and ethnicity data) may have been collected by projects from teachers, in the form of their assessment of student identity, and not directly from high school students (via student surveys) themselves. Given these inconsistencies and efforts to move the CS10K community toward district RRB/IRB compliance in the future (and further reasons discussed in later sections of the report), only the total number of students reported as reached/enrolled in CS10K teacher classes will be included in this analysis.

FINDINGS

How many teachers have been trained in CS10K projects? (Question 2)

The 20 responding projects reported they provided instruction to 683 teachers new to CS10K from periods 2011-2012 to 2014-2015. Figure 2 shows that 56 teachers participated in PL during the 2011-2012 period, increasing to 357 teachers in the period 2014-2015.

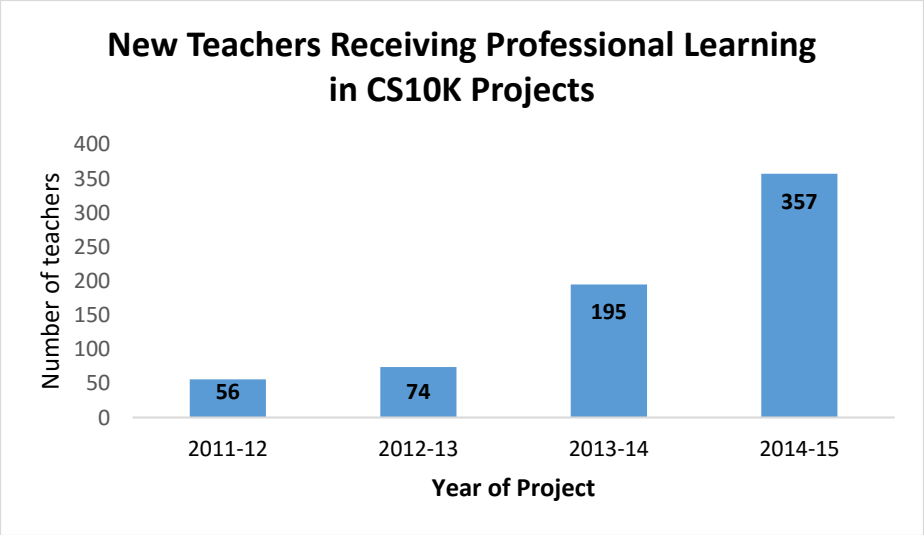

Figure 2. New CS10K teachers receiving PL during each reporting period

Figure 3 shows that programs provide instruction to an average of 14 new to CS10K teachers per year in the first year and to almost twice that many by the third year. Interestingly, projects vary in their capacity to provide PL. Figure 3 also shows the range of teacher participants with a maximum number of (maximum = 43) three times greater than the average (*m*=14) for the first yea, and similarly for the second year. This range does not persist into the third year, but it does indicate that some projects are reaching many more teachers than others. Figure 4 suggests that the number of teachers served by a project tends to increase, on average, over time and that there is a high amount of variation in number of teachers per project in any given year.

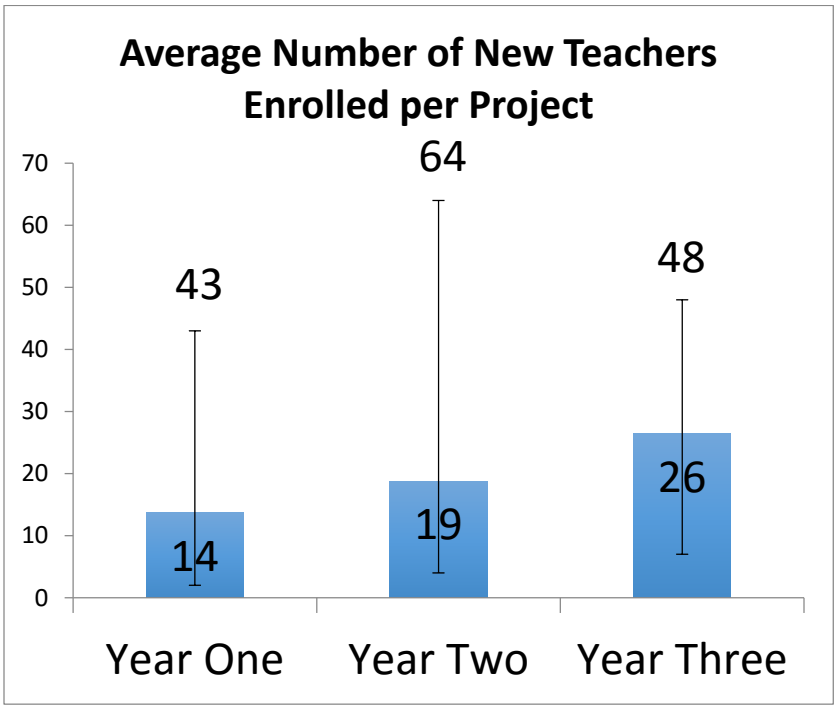

Figure 3. Average number of new CS10K teachers by project year

Figure 4 provides a more detailed breakdown of the number of teachers per project, by year (divided by ECS versus CSP-focused projects). This figure suggests the same trend as figure 3, above, but that there is a somewhat larger amount of variation among ECS projects, which tend to exceed the enrollment of CSP projects in their third year.

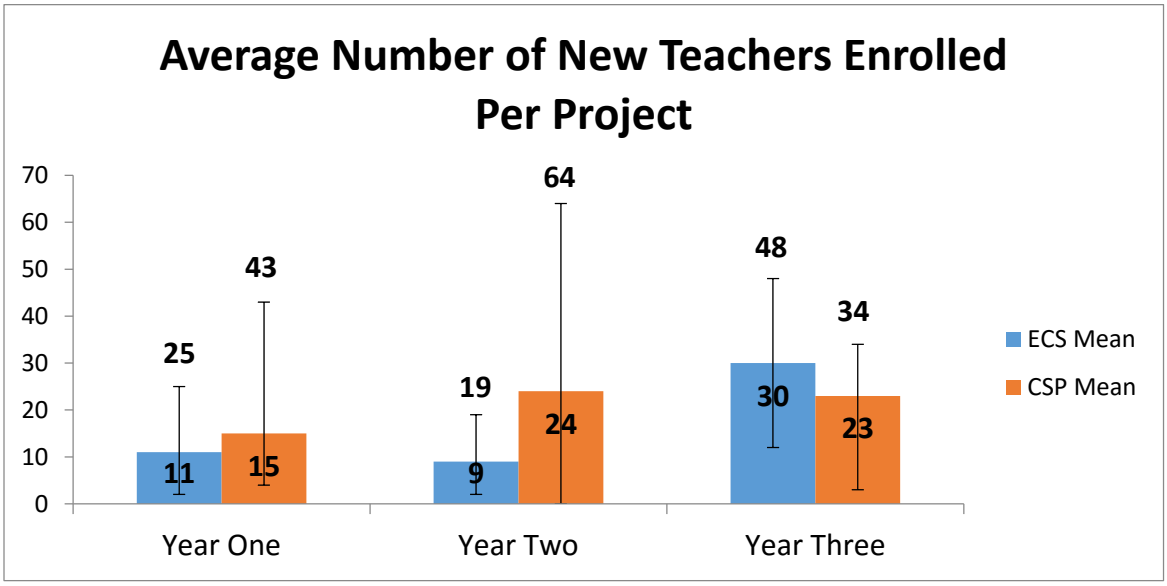

Figure 4: Average number of new teachers enrolled by project year and by project type

Teaching Experience

The new to CS10K teachers are overall an experienced group of K-12 and to a lesser extent CS teachers. Of the projects reporting, there are #% of teachers with more than six years of K-12 experience (Figure 5) and #% of teachers with more than six years of CS experience (Figure 6) (Although the EWG provided a definition of CS "as the study of computers and algorithmic processes including their principles, their hardware and software design, their applications, and their impact on society" teachers were asked to self-report their experience and may have interpreted other courses, such as keyboarding as CS).

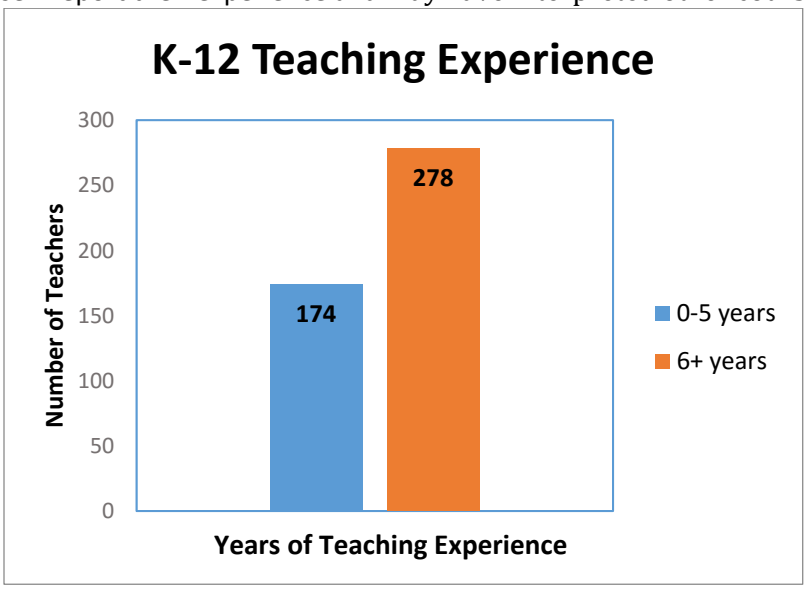

Figure 5: Years of CS teaching experience among CS10K PL participants

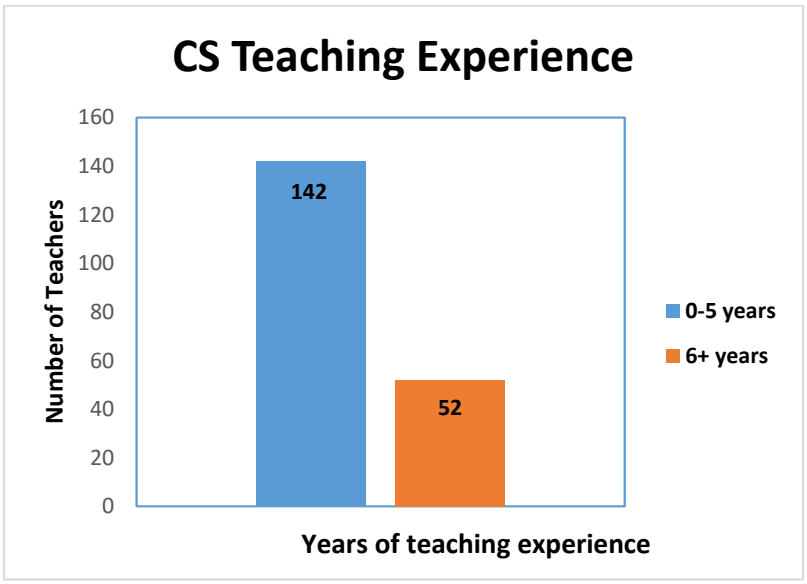

Figure 6. Years of CS teaching experience among CS10K PL participants

While the teachers are experience overall, there is a pattern of teaching experience is emerging across projects. CS10K projects generally offer instruction to a new group of teachers each year for three years. In the first year of program implementation, they serve teachers with more K-12 and more CS teaching experience (see figures 7 and 8) than in later PL years.

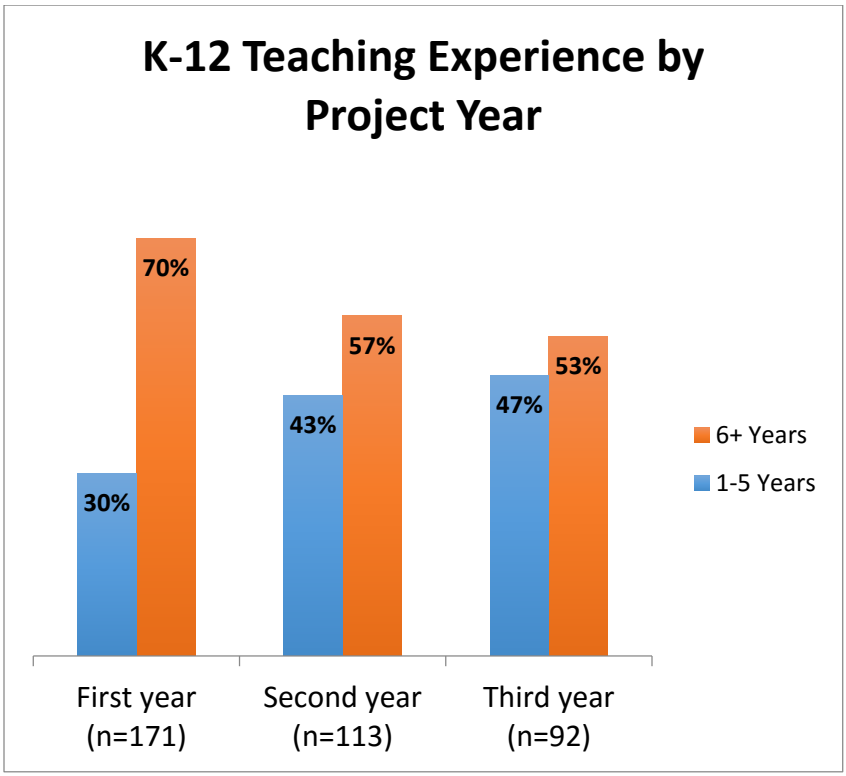

Figure 7. General K-12 teaching experience by project implementation year

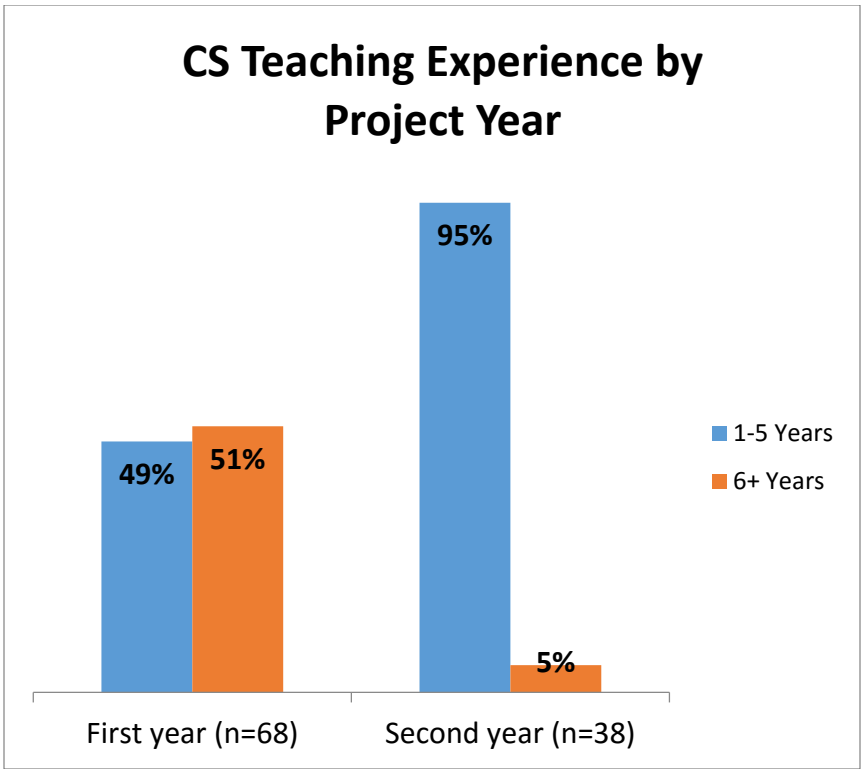

Figure 8. Self-reported CS teaching experience by project implementation year[5]

The CS experience, however, drops significantly during the second year of the project. The EWG hypothesizes that the projects recruit "master" teachers for the first year of implementation (perhaps using master teachers to help design curriculum, pilot curriculum, or perhaps the master teachers are more easily identified and recruited). The second year, the projects are reaching novice CS teachers.

How diverse are the teachers that have been trained through the CS10K program? (Question 6)

Gender

Among the new CS10K teachers for whom we have data, we see an even mix of male and female participants (Figure 9).

[5] Year 3 data was largely unavailable and there are not enough cases to report a meaningful percentage.

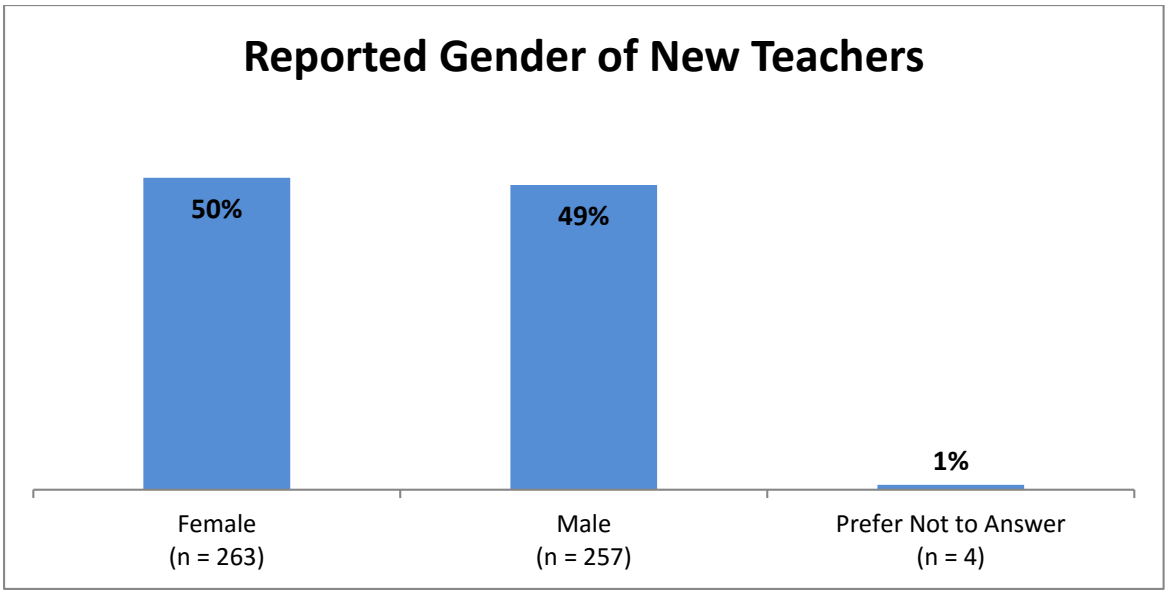

Figure 9. Gender breakdown of new CS10K teachers across all projects and years (N = 524; data missing or unavailable for 158 cases)

Race and Ethnicity

Instructors who participated in CS10K projects identified primarily as White and non-Hispanic. This proportion was reasonably consistent across projects reporting data (see Figures 10 and 11).

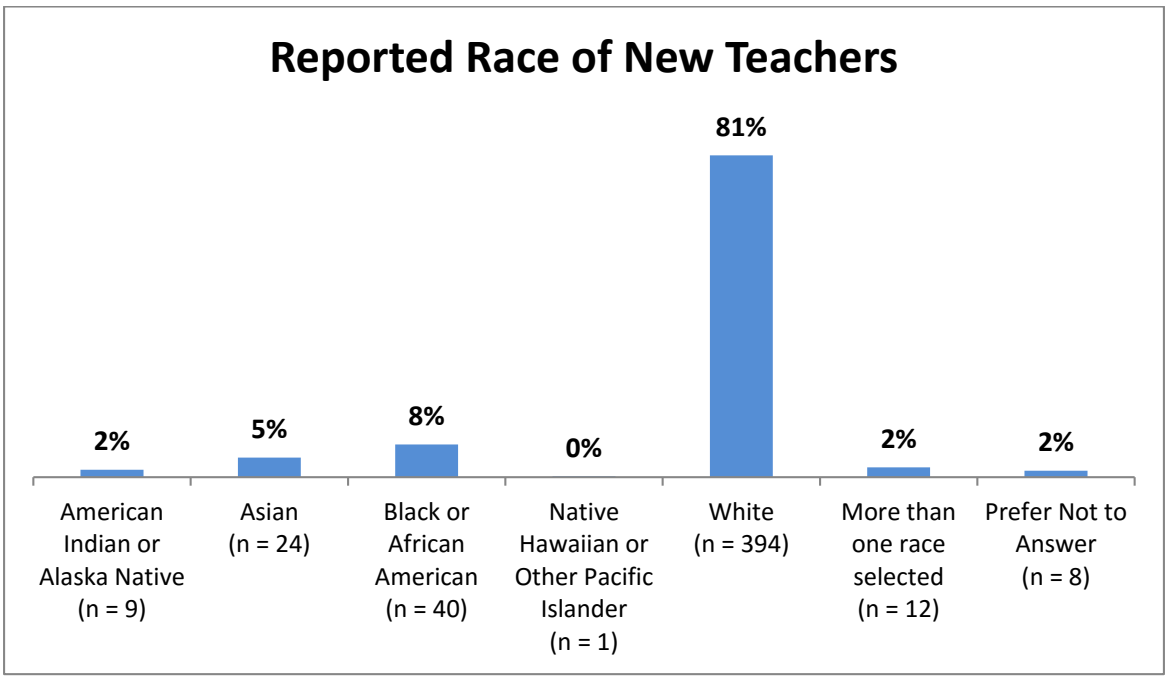

Figure 10. Racial demographics of new CS10K teachers across all projects and years (N = 488; data missing or unavailable for 194 cases)

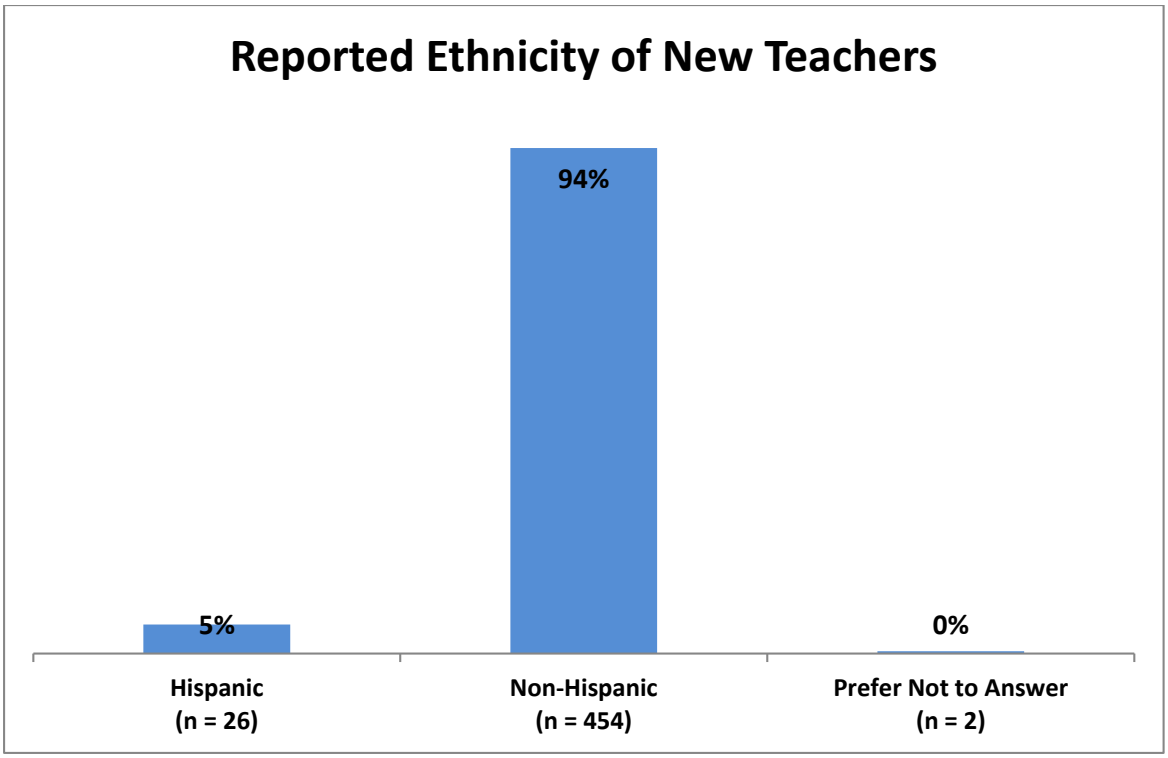

Figure 11. Ethnicity of new CS10K teachers across all projects and years (N = 482; data missing or unavailable for 200 cases)

Disability Status

We are currently working to collect disability data for participating teachers. However, in our current data set, the data are largely incomplete for most years. Collecting data on students with disabilities has been a challenge for many projects and a specific area of focus for the EWG. Two views guide our work: 1) Richard Ladner of the CS10K project ACCESS Computing recommends that not collecting data on students with disabilities only serves to further marginalize this population; and 2) the federal government and K-12 institutions place heavy restrictions on these data to protect students. The CS10K evaluation community will need further guidance (see Implications of Accessing K-12 Student Level Data below) on appropriately collecting these data so that we may adhere to both guidelines.

How many CS10K teachers are teaching with the instructional materials/approaches/curricula on which they have been trained to teach? (Questions 3 & 4)

This section answers two questions: 1) How many CS10K teachers are teaching with the instructional materials/approaches/curricula on which they have been trained to teach in PL; and 2) How many teachers have been through CS10K PL but were *not* teaching with the instructional materials/approaches/curriculum on which they were trained to teach at the time of data collection?

We asked evaluators to gather information about the extent to which teachers were utilizing the professional development materials within the class they were teaching for which it was intended (in most cases CSP or ECS). The intention of this question was to gain a better understanding of the status of implementation of the ECS and CSP materials by CS10K teachers. Most projects did not collect these data in previous years; therefore, the figures reported are for the most recent year only (see Figure 12), and only for those projects that returned data to us in the standardized data shell format. In addition,

these data were largely not collected among CS10K projects. However, of the 84 teachers trained in projects reporting this data, 38% teach primarily using the materials/approaches/curricula from their professional learning experiences along with supplementary materials, and 33% use none of the materials from their experience. We know that for at least one project (IDOcode) the teachers don't plan to implement until the 2015-16 academic year which may be heavily skewing the data.

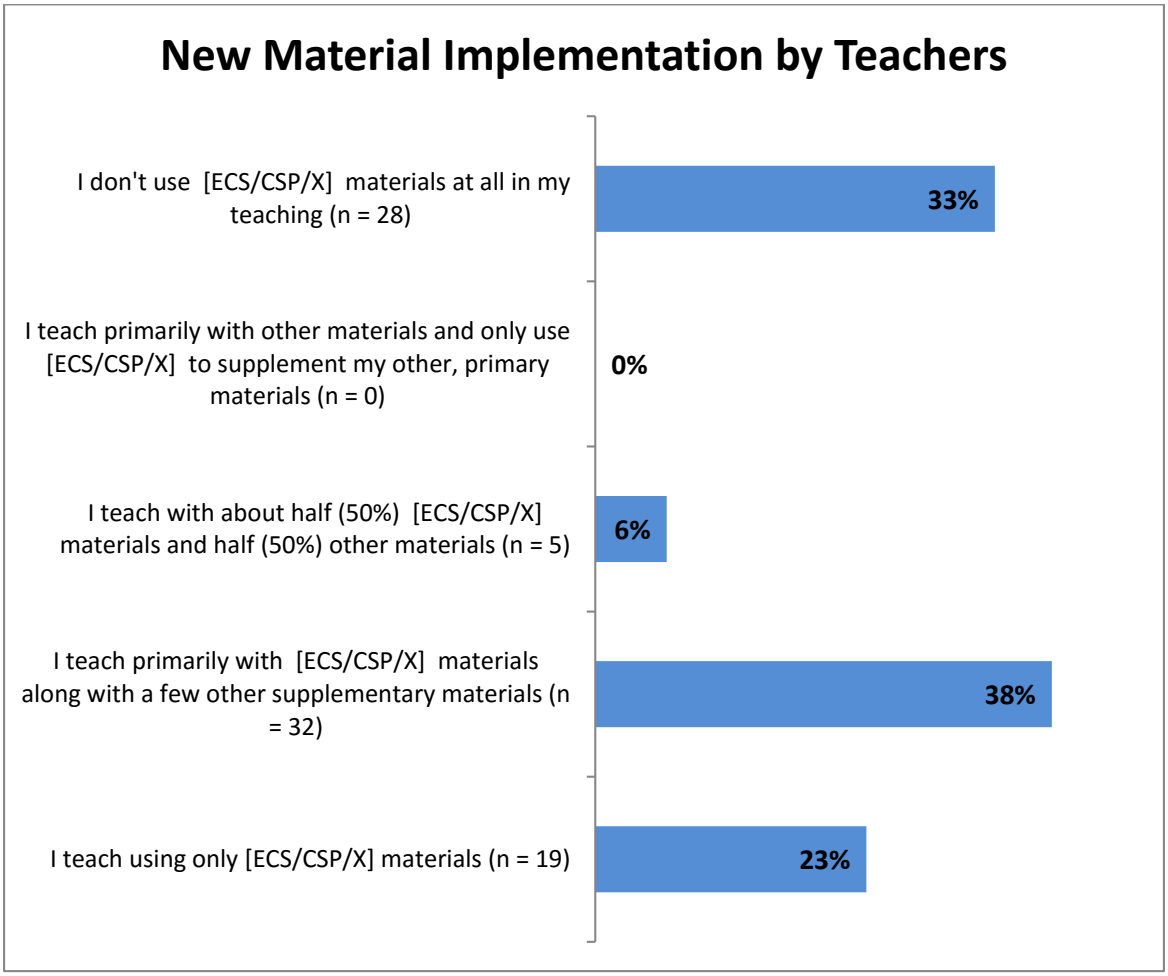

Figure 4. Reported use of materials/approaches/curricula among 2014-2015 CS10K teacher participants (N = 84; data missing or unavailable from 273 cases)

How many CS10K high schools are there? (Question 5)

The following map summarizes the locations of the public schools from which teachers from CS10K projects (which submitted usable school location data) teach. Each dot represents a single high school. The EWG collected school names and addresses from CS10K projects using the common elements data shell and verified each school against the National Center for Education Statistics (NCES) Elementary and Secondary Data System (ELSI). Overall, 463 unique public schools could be identified and found through the ELSI database in 35 states plus the District of Columbia. Private schools were not included because the data is inconsistently reported.

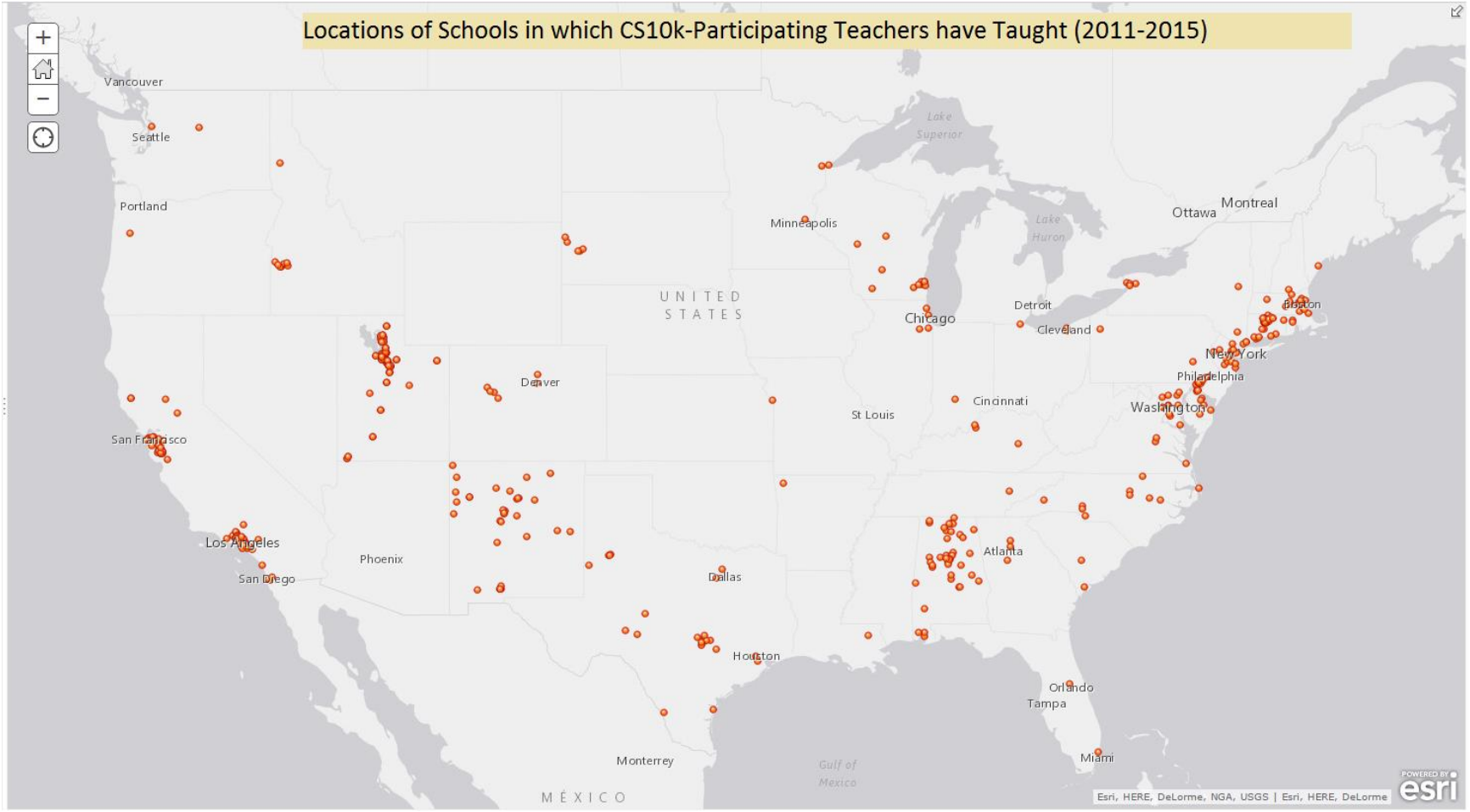

Figure 13: Map of locations of public schools from reporting CS10K projects

While 463 schools were verified against the ELSi database, approximately 630 schools were reported across 20 of the projects submitting data. An accurate number is not possible because we could not determine whether some school names were duplicates. There were a variety of reasons why some schools could not be identified, most prominently as follows:

- The search filters used in the NCES ELSi system were constrained to only look for public schools, private schools were excluded
- Some schools were not named properly, or were simply not present in the NCES ELSi system, and thus could not be found
- Some schools were not accompanied by enough background information (e.g. full name, zip code) to identify a single, positive match

How many and how diverse are the high school students reached through the CS10K program? (Question 7)

Each of the 463 public schools depicted in the above map was examined to determine basic demographic information about its student body, including race, gender and socio-economic status. This information was then compared to the same characteristics of a nationwide sample of public school students in the continental U.S. to determine the extent to which the body of students who were potentially reached through CS10K initiatives were representative of American public school students in general. The results can be seen in Table 5, below. The CS10K schools are largely representative of schools across the nation with a few exceptions: CS10K schools tend to have more students identifying as Asian/Pacific Islander, Hispanic, and White and fewer students identifying as Black/African American.

CS has a history of being found in schools with a high socio-economic status (SES) student body and absent in low-SES schools and many of the CS10K projects are trying to reach underrepresented schools and students. Our analysis indicates that schools offering computer science are seemingly diverse and in many cases as diverse as the national population of students. However, Margolis (2008)[6] cautions that school diversity may not extend to the computer science classroom and that numerous factors may limit or prevent access altogether. In this light, a promising indicator is that though the difference in the percentage of students eligible for free or reduced lunch may seem stark (51% in public schools nationwide vs. 41% in CS10K public schools) it is possible it is less than it would have been without CS10K.

[6] Jane Margolis. 2008. *Stuck in the Shallow End: Education, Race, and Computing*. The MIT Press.

Table 5: Demographic breakdown of public schools students at CS10K project schools versus national public school students (continental U.S. only)

	National Public School Students (N = 49,157,787)	CS10K Public School Students (N = 565,310)
% American Indian / Alaska Native	1%	1%
% Asian or Asian / Pacific Islander	5%	8%
% Hispanic	24%	26%
% Black/African American Students	16%	13%
% White students	51%	49%
% Hawaiian National / Pacific Islander	<1%	<1%
% Two or more races	3%	2%
% Male	51%	51%
% Female	49%	49%
%Free or reduced lunch eligible	51%	41%

As another measure of the representativeness of CS10K project schools, the schools compared in the table above were examined in terms of their "Urbanicity."[7] The comparison of CS10K public high schools versus all public high schools in the continental U.S. can be seen in Table 6, below. In general, CS10K schools are fairly well in-line with public schools across the nation with two exceptions:

1. CS10K public schools are noticeably more likely to be in a large suburban area
2. CS10K public schools are noticeably less likely to be in distant rural location

[7] NCES describes urbanicity as the extent to which a school is "urban-centric," which is a composite of the location (i.e. city, suburb, town or rural locale) and the relative population of the area

Table 6: "Urbanicity" of CS10K project high school locations versus national public high schools in continental U.S.

	National Public Schools (N = 98,547)	CS10K Public Schools (N = 463[8])
City: Large	14%	14%
City: Mid-size	6%	9%
City: Small	7%	10%
Suburb: Large	26%	37%
Suburb: Mid-size	3%	2%
Suburb: Small	2%	2%
Town: Fringe	3%	3%
Town: Distant	6%	4%
Town: Remote	5%	4%
Rural: Fringe	11%	8%
Rural: Distant	11%	2%
Rural: Remote	6%	3%

SUMMARY AND RECOMMENDATIONS

Lessons Learned from the Data

CS Teaching/Education Overall

The EWG is not yet in a position to report on the state of CS teaching and CS education overall. This is an important task to pursue and will continue into the second year of the effort. Specifically, we will work with other groups (CSTA, NCWIT etc.) to try to develop an approach to ascertain the number of CS teachers in the United States, although this is a challenging task as the ways the people define CS is inconsistent. In collaboration with CS10K projects and through participation in the greater CS education community, we believe CS is on the cusp of entering the educational cannon as a core discipline (like math or English/Language Arts) and that any investment in evaluating the field is important.

Further, data for this report and our conversations with evaluators and PIs shows that not all teachers who participate in CS10K PL end up teaching the course for which they were trained. As a community, we need to better understand the barriers (environmental, organizational, and personal) and how to reduce them.

[8] Number of high schools identified in ipeds

Lessons Learned About Collecting the Data

Engaging the Community of Evaluators

Counting the participants in a project through evaluator and PI reports requires a strong level of buy-in. The formation of the EWG was a first step in creating support. EWG members are all part of the evaluation community. As such, they know the evaluators, appreciate the resource constraints, and are tasked with same burden being asked of their peers. The data collection tools and processes were designed with a practitioner perspective and although some items requested were aspirational, most were reasonable to request and straightforward to provide.

Though the design of the data collection tools and process was bottom-up, the top-down request from NSF was an essential part of establishing the authority for the EWG to lead the effort. NSF has historically invested in developing a community of evaluators for the CE21 programs. Thus the CS10K community of evaluators is a particularly strong with many having worked together for years at NSF-sponsored meetings and as pioneers in CS education evaluation space and were happy to support the request for common data.

The blend of virtual, face-to-face and email/phone interactions allowed the EWG to reach almost all members of the CS10K evaluator community. The phone and email conversations between the liaisons and the projects allowed for customized support and accountability while the larger webinar and meeting at the PI conference legitimized the request and sent the message that "we're in this together." The EWG also timed the communication so that by the time the community gathered in person they had already interacted through the webinar and with their liaison.

Finally, although the effort targeted evaluators, the PIs were included in all webinars, meetings and emails as to generate awareness and support for the evaluator.

Implications of Accessing K-12 Student-level Data

Background information about students enrolled in ECS and CSP courses is important. The NSF CS10K community wants to know details about the students in schools involved in these courses (e.g., race, ethnicity, gender, and disability/IEP, or "Individualized Education Program" status). But schools and school districts are extremely complex systems, and they have established regulations (rooted in federal and local confidentiality laws) about what can be accessed, by whom, and for what purposes.

Any research in public school districts (during contract hours and/or on district property and/or with minors) requires some level of district-level approval. To work in school districts, researchers must first submit plans to their internal Institutional Review Board (IRB) and second, submit a request to school district IRB departments or Research Review Boards (RRBs). Information about how to request permission to conduct research is more easily accessible for some school districts than others. Some districts have explicit, published research request procedures and review board timelines, others need to be contacted directly to find out what is needed to receive approval to do data collection, or request access to data from districts. These requirements apply not only to minors, but also to any subject in the school (i.e., teachers and principals). This is particularly important if data about minors are collected to share with a wider research/evaluation audience. Classroom teachers, as they are not researchers, may not be aware of what types of student-level data they are legally restricted from providing to third party researchers. Thus while they may have access to district information about students, district regulations prohibit them from sharing that information with others. Researchers must take responsibility for being

versed in the district requirements to ensure they do not ask teachers for information that the district does not actually permit them to share (but that they may not be aware of as non-researchers).

When doing research in school districts, following proper channels to receive approval from district IRB/RRBs to collect or obtain any data in schools or from districts can be a challenge, for a number of reasons:

- District requirements around conducting research and requesting student data vary. Every district is different and has its own set of requirements and format for how those requirements are delivered to it for approval.
- Most districts stress that research and data requests *must come through proper channels at the district level*. Once approval is received, only then can requests to the district be made for student level data, and contact be made with schools to request school leader permission to begin any direct data collection (e.g., questionnaires or interviews with students and teachers outside of weekend PL sessions). There may be some cases where work in only one or two schools does not require district RRB approval (or where an evaluation project is considered a "district initiative"), but the district should always be contacted to ask.
- Many districts require "active parental consent" for any participation of minors in research that is not considered a "district initiative" (i.e., with active consent, parents must sign and return a consent form indicating their child has permission to participate in any data collection. "Passive consent," in comparison, requires a signed and return consent form only when do *not* grant permission for student participation. If no consent form is returned to a teacher, permission is assumed to be granted), Types of data collection in districts that require any type of parental consent can include an online questionnaire, interviews, focus groups, submission of examples of classroom work, and in some cases, requests for student level data from the district). Active consent requires researchers to ask teachers whether they are willing to have their students participate in research activities; provide teachers with ample copies of active parental consent forms to distribute to students; hope that parents read and sign the forms indicating students have permission to participate; hope that students return signed forms to the teacher; and go to schools to pick up forms, or ask teachers to send the signed forms – i.e., the forms for students who DO have permission to participate in research – back to researchers. Only students who returned signed forms indicating they *do* have permission to participate in research may be involved. This is a burdensome process for all involved, and significant portions of teachers, parents, and students do not complete all required steps.
- Some student data are considered more sensitive than others (e.g., race/ethnicity, disability status/IEP). Most districts have rules about only providing data in aggregate so it is not personally identifiable. There are often "minimum number of student" stipulations, too, so that in situations where there are 10 or fewer students of interest, districts may not allow researchers access to the information for fear of confidentiality breach.
- Some data requests to districts have associated fees. That is, the district may need to pay their data specialists by the hour (possibly at a rate of $100/hour) to pull and sort the requested data (e.g., Denver, Washington DC, etc.).
- In general, the district IRB/RRB process requires considerable time, especially if they have a set application. Evaluators must factor time into their research plans and contracts with project PIs, and be certain PIs understand that there can be long waiting periods to receive approval from school districts (some have set research review dates that only happen a few times a year). In some cases, this can actually take months and requires official agreements or memoranda of agreements.

- Districts often ask researchers to justify the time and effort required for data collection in schools by including a section in the application asking how the research benefits the school district. This helps them sort through the many research requests they receive each year.
- Researchers/evaluators working on projects that include teachers and/or students from multiple school districts are tasked with going through this time-consuming process with each district.

Contacts in district IRB/RRB departments/divisions may carry a range of titles. In most cases, classroom teachers do not know who in their district handles external research/data requests and these people have a wide range of titles.

Districts may decide to grant or deny approval for a specific research request to their district for a number of reasons. For example, if a district already has many external research requests, they may decide the burden placed on teachers, students, or their own data department is not worth the benefit of the research to their district. Moreover, they may simply have too many internal research projects in progress to add more in an academic year. Overall, CS10K projects face the possibility that a school district might agree to partner on a project and refuse to allow data collection on their teachers and students.

Lessons Learned about the Context

Understanding the context of CS10K projects provides a greater understanding of the data. The first round of data collection left the EWG with several questions, which will be addressed in year 2. Specifically we wish to:

- Determine whether or not teachers are currently teaching the CS course for which they were prepared, and/or
- Describe how the PL is being used and the extent to which the PL experience is implemented. For example, are approaches/techniques/strategies learned in the PL being used in courses a teacher teaches outside of their ECS or CSP class?

We have also learned that we need to better align data collection dates with teaching schedules for district-level data like student participation. Typically, district-level data are available by August, requiring a shift in collection schedule to asking for completed data shells in late August/September.

The data collection has raised more questions than it has answers. For this reason, input from NSF on desired items is critical. For example, does NSF want to know about the specific CS background teaching experience of CS10K teachers?

Finally, we recommend that NSF require reporting of basic information (described above) as a condition of the project award and included this requirement in the program solicitation. Though there is enormous good-will within the evaluator community, the task is burdensome. We further recommend the development of a reporting system (possibly looking to the Math Science Partnership for guidance). Requiring reporting through a dedicated system will streamline the process and improve the quality and quantity of data provided.

PLAN FOR 2015-2016

The EWG has been tasked with the following:

1. Refine the common elements data shell that contains demographic items for teachers and students participating in or affected by CS10K projects. Members of the EWG and CS10K evaluators will continue to refine a second version of the data shell.
2. Continuously improve the data collection process to ensure accurate and timely data from CS10K projects.
3. Collect data for the 2015-16 academic year (July 1, 2015 - June 30, 2016).
4. Curate and begin answering research questions from the broader research and evaluation community using the submitted data.
5. Coordinate with other CS teacher training programs (Code.org, Project Lead The Way and other STEM+C initiatives providing CS professional learning) to explore the broader landscape of CS teacher preparation in the United States.
    a) Identify and share among the broader CS evaluation community instruments, scales, constructs, and approaches.
    b) Form a set of recommendations for ongoing data collection among evaluators.
    c) Meet annually with evaluators at NSF-funded meetings of CS10K PIs and Evaluators to provide a snapshot report of CS10K teachers based on data the programs provided in 2014-15 and to share instruments, methods, and promising directions. This will also serve as an "induction" for evaluators new to CS and invite them to become active participants in the CS evaluation community.

Relevant to the second round of data collection, the team will be refining the data collection tools and process. Approximately 12 new projects will begin in September 2015, and those projects will need to know what must be measured and the set of recommended scales and constructs that have been used with successful projects in the past.

Tools

This fall the EWG will interview sites that submitted information to the EWG to validate the processes and tools put in place. We want to be clear of the interpretation of items and the feasibility of implementation, and continue building a culture of common data collection. From these interviews, we will modify the common elements data shell and technical documents. The EWG also requests that NSF identify the fields most critical for their reporting. This will help minimize the burden on evaluators while maximizing the value to the program.

Unlike the first round of data collection, the next phase of data collection will limit data collection to the current, 2015-16 school year. we also need to adjust the timing of the data collection since many projects are administering end-of-year surveys in June and do not have results until late July. We ask that CS10K projects follow district guidelines and acknowledge that data procured district offices may not be available until mid-August. To date, we know the following fields will be changed:

1. Teacher demographics: Will we adjust the "race" category and eliminate "more than one race selected" so that evaluators can report exactly what teachers reported (i.e. multiple options selected, or an individual choice selected. This may be changed to match the new Census

Bureau guidelines (rather than OMB guidelines). We will move the fields about use of PL material to the teacher experience tab.

2. Teacher experience: We will include the questions about K-12 and CS experience. We will modify the tab for greater clarity of questions.
3. Classroom-level data*:* only number of students taught will be collected. We will provide additional support to evaluators to further the ability to work with district offices to request student level data. We will also rename this tab "student-level data"
4. School data: We will consider adding an option for "public" or "private" school to facilitate mapping.

The technical document will be updated to match revisions to the data shell.

The EWG will explore alternative ways of collecting the data. The excel spreadsheets worked well, but a database system would be easier for entry and analysis, improve accuracy, and facilitate analysis. We will determine the feasibility of creating such a database.

Process

Overall, the EWG is pleased with the data collection process and is recommending minor changes for the next data collection cycle.

Webinar: We will hold another webinar in October 2015 as an orientation for the projects funded in 2014 and 2015, though we will invite all evaluators and PIs from the community to participate. This proved to be an effective way to communicate with the evaluators and PIs who chose to join and helped establish a community last year. The 2015 webinar will present the findings to date and provide an orientation to the revised data shell. We strongly encourage NSF to promote the webinar, particularly to the 2014 funded projects, many of which will be working with teachers implementing ECS or CSP for the first time, and for the 2015 funded projects to bring them into the community and help them prepare their evaluation plans and instruments.

CE21 PI Meeting 2016: At this meeting we plan to hold another workshop for evaluators and PIs to discuss the data findings to date and to provide implementation support. There will be a second meeting for evaluators to discuss the broader topics in evaluating CS education.

Liaisons: The first year data collection was successful because of the liaison system set up between the EWG and projects. This system began to generate evaluator buy-in, provided support to project evaluators working to provide the EWG with requested data, and outlined expectations moving forward. It also provided the EWG with a detailed understanding of each project's circumstance. The second year of data collection will be modified as we have already created a strong base for the common data collection. In year two, we will continue the liaison model, but with a focus on supporting the new projects prepare for data collection. Liaisons will also connect with projects to understand any unique challenges and submission deadlines.

Once submitted, data will be reviewed centrally (SageFox) and any specific follow-up questions will be handled by the liaison.

CS10K Community Evaluators Group: We will continue to make all our materials available to project evaluators on the CS10K Evaluators Group in the CS10K Community, including the webinar recording for

those unable to attend. We will encourage project evaluators to share and discuss their evaluation materials and experiences in the group.

APPENDICES AND DATA TABLES

Appendix A: Data tables

i) Aggregated tables
ii) Returning teacher's data

Appendix A: Data Tables

Aggregated Teacher Descriptives

		2011-2012		2012-2013		2013-2014		2014-2015	
	Number of Projects Providing Data	3		5		11		17	
		New	Returning	New	Returning	New	Returning	New	Returning
	Total Number of Teachers*	50	6	74	14	195	29	357	99
How are teachers teaching with PD materials/approaches/and curriculum?	I teach using only [ECS/CSP/X] materials	0	0	0	0	4	2	19	8
	I teach primarily with [ECS/CSP/X] materials along with a few other supplementary materials	0	0	0	0	15	0	32	10
	I teach with about half (50%) [ECS/CSP/X] materials and half (50%) other materials	0	0	0	0	2	0	5	0
	I teach primarily with other materials and only use [ECS/CSP/X] to supplement my other, primary materials	0	0	0	0	1	0	0	0
	I don't use [ECS/CSP/X] materials at all in my teaching	0	0	0	0	5	0	28	0
	Data Not Available	50	6	74	14	125	27	220	75
Gender	Female	8	2	2	5	82	8	169	47
	Male	10	0	8	6	80	7	159	47
	Prefer Not to Answer	0	0	0	0	0	0	4	0
	Data Not Available	32	4	64	3	33	14	23	5
Ethnicity (Hispanic)	Yes	0	1	0	0	6	0	19	3
	No	14	1	9	11	148	15	282	52
	Prefer Not to Answer	0	0	0	0	0	0	2	0
	Data Not Available	36	4	65	3	41	14	31	16
Race	American Indian or Alaska Native	0	0	0	0	4	0	5	4
	Asian	2	1	2	2	8	4	11	4
	Black or African American	1	0	0	1	15	1	24	8
	Native Hawaiian or Other Pacific Islander	0	0	0	0	0	0	1	0
	White	10	0	6	8	130	9	248	57
	More than one race selected	0	0	1	0	4	0	7	2
	Prefer Not to Answer	1	0	1	0	1	1	5	0
	Data Not Available	36	5	65	3	36	14	29	14
Disability (Y/N)	Yes	0	0	4	0	8	0	12	4
	No	4	0	0	3	90	4	171	49
	Prefer Not to Answer	0	0	1	0	1	0	4	0
	Data Not Available	46	6	68	11	96	25	170	45

*Note that, due to missing data on some submitted sheets, the total number of responses provided for any given question may not sum to the total number of new or returning teachers for a given year

Teacher Experience Aggregate

		2011-2012		2012-2013		2013-2014		2014-2015	
	Number of Projects Providing Data*	3		5		11		17	
		New	Returning	New	Returning	New	Returning	New	Returning
	Total Number of Teachers	50	6	74	14	195	29	357	99
Number who replied that they had taught K-12 for the given number of years, including the current year	1 Year	0	0	0	0	14	0	37	2
	2-3 Years	0	0	2	0	16	1	44	6
	4-5 Years	0	0	2	0	18	0	41	5
	6-10 Years	2	0	2	0	26	1	66	13
	11-15 Years	0	0	0	1	26	1	46	12
	16-20 Years	1	0	0	1	15	1	29	9
	21+ Years	6	0	1	2	28	4	30	18
	# Years Unknown	1	0	0	0	1	0	14	0
	Data Not Available	40	6	67	10	51	21	53	33
Number who replied that they had taught computer science for the given number of years, including the current year	1 Year	0	0	1	0	7	0	84	0
	2-3 Years	0	0	1	0	2	0	25	3
	4-5 Years	0	0	1	0	9	0	12	4
	6-10 Years	0	0	1	0	2	0	25	8
	11-15 Years	0	0	0	0	3	0	3	1
	16-20 Years	0	0	0	0	2	0	4	3
	21+ Years	0	0	1	0	3	1	8	0
	# Years Unknown	10	0	0	4	41	5	14	33
	Data Not Available	40	6	69	10	72	23	99	32
Exploring Computer Science	Not at All	4	0	0	3	25	4	49	23
	1 Year	0	0	0	0	2	0	5	0
	2-3 Years	0	0	0	0	0	0	0	4
	4-5 Years	0	0	0	0	1	0	0	0
	6+ Years	0	0	0	0	0	0	0	0
	# Years Unknown	1	0	0	1	15	1	2	8
	Data Not Available	45	6	68	10	107	23	232	48

		2011-2012		2012-2013		2013-2014		2014-2015	
	Number of Projects Providing Data*	3		5		11		17	
		New	Returning	New	Returning	New	Returning	New	Returning
	Total Number of Teachers	50	6	74	14	195	29	357	99
AP Computer Science Principles	Not at All	0	0	0	0	4	0	48	3
	1 Year	10	0	0	0	43	0	33	6
	2-3 Years	0	0	0	4	0	5	2	40
	4-5 Years	0	0	0	0	4	0	0	0
	6+ Years	0	0	0	0	0	0	1	0
	# Years Unknown	0	0	0	0	0	0	3	0
	Data Not Available	40	6	68	10	90	23	202	34
AP Computer Science A	Not at All	0	0	0	0	13	0	41	10
	1 Year	0	0	0	0	2	0	1	0
	2-3 Years	0	0	0	0	0	0	0	0
	4-5 Years	0	0	0	0	1	0	6	1
	6+ Years	0	0	0	0	0	0	6	2
	# Years Unknown	5	0	0	4	23	5	0	22
	Data Not Available	45	6	68	10	101	23	232	49
Microsoft Office suite or keyboarding	Not at All	0	0	0	0	2	0	23	2
	1 Year	0	0	0	0	2	0	1	0
	2-5 Years	0	0	0	0	1	0	8	1
	6+ Years	0	0	0	0	2	0	11	1
	# Years Unknown	0	0	0	0	0	0	0	0
	Data Not Available	50	6	68	14	130	28	232	79
Other computing courses	Not at All	0	0	0	0	3	0	25	4
	1 Year	0	0	0	0	1	0	4	0
	2-5 Years	0	0	0	0	2	0	8	2
	6+ Years	0	0	0	0	3	0	11	0
	# Years Unknown	5	0	0	4	36	5	5	29
	Data Not Available	45	6	68	10	85	23	211	28

*Note that, due to missing data on some submitted sheets, the total number of responses provided for any given question may not sum to the total number of new or returning teachers for a given year

Student Participation Aggregate

	2011-2012		2012-2013		2013-2014		2014-2015	
Number Providing Data	3		4		11		11	
Total number of teachers reporting student data	50		50		147		226	
Total number of course sections taught by project teachers	67		57		105		145	
Total number of course sections for which student data are reported	0		0		25		97	
	Enrolled at beginning	Passed	Enrolled at beginning	Passed	Enrolled at beginning	Passed	Enrolled at beginning	Passed
Total Number of Students	2746	303	2542	0	5786	1026	7311	2687