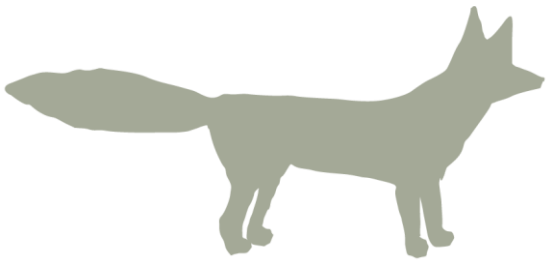
CS10K COMMON DATA REPORT

Evaluation Working Group
Common Data Collection Report

2017-2018
Award #1228355



SAGEFOX CONSULTING GROUP

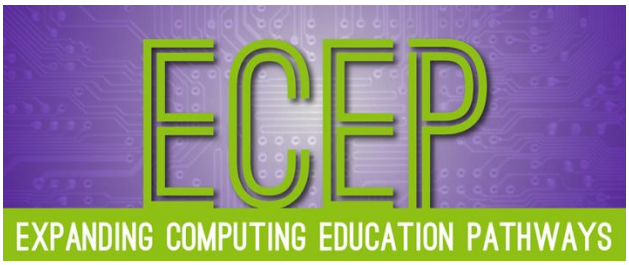


Table of Contents

Table of Contents 1
Chapter 1: Introduction, Mission and Context 2
Evaluation Working Group Guiding Questions 3
The Computer Science Education Context 4
Chapter 2: Findings 8
Methods and Limitations 8
Caveats 10
Findings 11
Teachers 11
Gender 13
Race and Ethnicity 13
Disability Status 14
Teaching Experience 15
Experience Teaching CS Courses 15
Students and Schools 20
Chapter 3: State Data 28
The Challenge 29
Findings 29
Data Analysis and Publication 31
Defining broadening participation goals 31
Preparing to collect statewide data 31
Data requests and lessons learned 32
Discussion 33
Next steps 34
Chapter 4: Lessons Learned and Recommendations 35
Lessons Learned 35
Recommendations 40
Appendix A: Memo: CS10K Close-Out Survey Feedback 42
Appendix B 48
Appendix C 51
Appendix D 51
Appendix E: Support Document Common data collection 55
Appendix F: Methods for the assessment of state data capabilities 59

Chapter 1: Introduction, Mission and Context

The National Science Foundation's (NSF) CS10K program (awards made 2012-2016) "aims to have rigorous, academic computing courses taught in 10,000 high schools by 10,000 well-prepared teachers. CS10K proposals have focused on high school computer science (CS) teachers, providing preservice and in-service teachers with courses, professional development opportunities, and long-term, ongoing support" (NSF Program Solicitation 15-537). This report is an effort towards answering three questions central to the National Science Foundation's CS10K program:

1. How many teachers have participated in professional development (PD) through CS10K-funded projects?
    a. What are the demographic characteristics of these teachers? And
    b. What is their teaching experience?
2. How many students did the CS10K projects reach in 2017-18?
    a. What are the characteristics of students that were reached through CS10K?
    b. What are the characteristics of the student subset who took the AP CSP exam?
3. How many schools have a trained CS teacher?
    a. What are the characteristics of the student body that has access to a course taught by a CS10K-trained teacher?

A fourth question is:

4. How feasible is program-wide data collection?
    a. What lessons can be learned through cross-project data collection.

The Evaluator Working Group (EWG), a group of Broadening Participation in Computing (BPC) and CS10K evaluators, is responsible for facilitating participation in a common data collection effort across CS10K projects that provide professional development (PD) to teachers in preparation for offering the Exploring Computer Science (ECS) course or the Computer Science Principles (CSP) AP exam course. This group successfully piloted the process in 2014-15 and conducted data collection in 2015-16, 2016-2017 and 2017-2018. The participants in the 2017-18 Evaluator Working Group include:

Rebecca Zarch	SageFox Consulting Group
Kathy Haynie	Haynie Research and Evaluation
Tom McKlin	The Findings Group
Christine Ong	UCLA, CRESST
Alan Peterfreund	SageFox Consulting Group
Stacey Sexton	SageFox Consulting Group
Gary Silverstein	Westat
Jeffrey Xavier	SageFox Consulting Group
Sarah Dunton	Expanding Computing Education Partnerships (ECEP)

Chapter 1 of this report describes the mission of the project and the computer science education context. Chapter 2 presents the data collected about teachers, students, and schools. This includes the number, demographic characteristics, and teaching experience of the teachers participating in CS10K-funded professional development; counts and demographics about the students reached and

students potentially reached by trained teachers. Appendix A provides a reflection on the cross-project data collection process. The third chapter documents what the EWG has learned about the potential value of defining a broadening participation goal and utilizing state-level data systems to track progress against this goal. The final chapter of the report presents lessons learned and recommendations.

Evaluation Working Group Guiding Questions

The EWG addresses nine critical questions that guide this report and future reporting of common data elements. Table 1 presents the questions along with potential data sources.

Table 1: Critical Questions

Q#	Common Question	Data Sources
Questions about teachers:		
Q1	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 ; EWG Common Elements Data Shell: Teacher descriptives/demographics
Q2	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
Q3	What fields are the CS10K teachers certified and/or teaching?	EWG Common Elements Data Shell: Teacher descriptives/demographics
Q4	How many CS10K teachers are teaching with the instructional materials/approaches/ curricula they have been trained to teach?	EWG Common Elements Data Shell: contact information, Teacher descriptives/demographics
Questions about the Students and Schools:		
Q5	How many CS10K high schools are there?	EWG Common Elements Data Shell: School data from evaluators
Q6	How many students have been reached through the CS10K projects?	EWG Common Elements Data Shell: student descriptives/demographics
Q7	How diverse (race/ ethnicity/ gender/ disability status/ CS experience) are the students who have been reached through the CS10K program?	EWG Common Elements Data Shell: Student demographics/District student level data, Teachers' schools
Q8	How many and how diverse (race/ethnicity/gender/disability status) are the high school students potentially reached through the CS10K program?	EWG Common Elements Data Shell: schools; NCES data
Q9	How many and how diverse (race/ethnicity/gender/disability status) are the high school students who took the AP CSP exam?	EWG Common Elements Data Shell: Student demographics

Our working group collectively judged these questions to be most critical with respect to capturing the impact across CS10K efforts. We acknowledge the variation among CS10K projects, most

notably the type of CS preparation offered through PD. This report is a product of the effort to answer these questions.

Participation in this data collection effort was voluntary. Though most projects submitted data, not all projects were able to provide data for each element, and it is unclear if the data provided represents 100% of the teachers trained, or a subset of whom participated in the project-level evaluation activities. As such, the findings presented in this report reflect the available data, which is incomplete. This is our "best snapshot" at this moment in time of CS10K projects but may not provide the full picture. The data also assume all CS10K projects are equally efficacious, as the EWG is not tasked with evaluating the relative strength of a particular project. Also, the relative size of the projects likely bias the data, particularly those with large cohorts of teachers, or those that are able to access large amounts of student data (typically through a district-level request).

This group successfully piloted the process in 2014-15 and conducted data collection in 2015-16 and 2016-2017 and 2017-2018.

In Years 3 and 4, we responded to the rapid shift in the national CS educational context (multiple PD providers; multiple funding mechanisms and general political support for CS) by exploring the need to access and potential value of state-level data systems to accurately measure the impact of teacher PD on CS education including identifying which courses are offered, at what schools, and which students are participating and succeeding in CS courses.

The Computer Science Education Context

At the start of the CS10K program in 2012, efforts to expand computer science education were sparse, with dedicated individuals taking advantage of opportunities to promote computer science, prepare teachers and develop curricular materials as they arose. Many of the early CS10K efforts involved Exploring Computer Science (ECS) rollouts in large cities such as Los Angeles, as well as the early field-testing of the CS Principles project in less than 100 high schools of colleges/universities nationwide. Since this project began in 2014 there have been substantial changes in the CS educational landscape leading to more strategic and large-scale planning for CS education for example:

- In January 2016, President Obama announced the CS for All initiative which greatly elevated the prominence of the work being done by many of the CS10K projects and others involved in CS Education[1].
- The development of the K-12 Framework for Computer Science[2] (released in 2016) with over 50 writers and advisors and hundreds of reviewers, leading to broad support and many states either adopting, or adapting, the framework.
- State education and policy leaders are creating policies to promote and expand CS education including the adoption of standards, development of curriculum, teacher professional development, and reviewing credentialing requirements. ECEP and Code.org have been at the forefront of supporting states' efforts, promoting and supporting the development of statewide strategic planning.

[1] https://obamaWhitehouse.archives.gov/blog/2016/01/30/computer-science-all
[2] https://k12cs.org/

- The Governors for CS initiative[3] has secured the commitment of 16 state governors to: 1) enable all high schools to offer at least one rigorous CS course; 2) fund professional learning opportunities so teachers can be prepared to teach CS content; and 3) create a set of high quality academic K-12 computer science standards to guide local implementation of CS courses.
- NSF has continued the commitment to CS education through the STEM+C[4] program and the 2017 debut of the CS for All: RPP Research Practice Partnership program[5] (which has absorbed the CS10K program). These, and other NSF programs are contributing to the preparation of high quality CS teachers across the nation.
- In fall 2017 the US Department of Education announced a $200 million investment into CS education, which will likely continue to advance the CS Education ecosystem in the nation.[6]
- An increasing number of organizations with a nationwide presence, such as Code.org[7], Bootstrap and Project Lead the Way[8] have been offering teacher PD, often to prepare teachers to teach ECS and/or CSP.

Along with these large-scale national efforts, NSF has focused resources and attention supporting the development of two important courses: Advanced Placement Computer Science Principles (CSP) and Exploring Computer Science (ECS) courses.

AP Computer Science Principles offers a multidisciplinary approach to teaching the underlying principles of computation. The course introduces students to the creative aspects of programming, abstractions, algorithms, large data sets, the Internet, cybersecurity concerns, and computing impacts. AP Computer Science Principles also gives students the opportunity to use current technologies to create computational artifacts for both self-expression and problem solving. Together, these aspects of the course make up a rigorous and rich curriculum that aims to broaden participation in computer science. The AP Computer Science Principles Curriculum Framework provides information about the Big Ideas and Computational Thinking Practices that are to be covered in the course, the two through-course assessments, and the end-of-year assessment. AP CS Principles was launched in September 2016 in over 2,700 schools. The largest AP rollout in history, 43,780 students completed the exam in May 2017; of this group, 30% were female students and 30% were underrepresented minority students. An estimated 70,000 students registered for the exam in May 2018[9]. There are currently a number of widely used CS Principles curricula, some of which have been endorsed by The College Board (e.g., Beauty and Joy of Computing, code.org). Starting in 2016-17, every course syllabi (i.e., at the individual teacher, school, or district level) is reviewed by The College Board prior to receiving AP course credit.

Exploring Computer Science (ECS) is an introductory CS course for high school students, which includes a year-long curriculum consisting of six units, approximately six weeks each, and an accompanying professional development program. ECS is designed to be implemented with high

[3] https://www.governorsforcs.org/
[4] https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505006
[5] https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf18537
[6] https://innovation.ed.gov/what-we-do/stem/computer-science/
[7] https://code.org/
http://www.bootstrapworld.org/index.shtml
[8] https://www.pltw.org/our-programs/pltw-computer-science
[9] https://www.totalregistration.net/AP-Exam-Registration-Service/AP-Exam-Score-Distributions.php

fidelity to the curriculum to ensure a cohesive scope and sequence and to support inclusive teaching practices. The course was first piloted by the NSF-funded Into the Loop project in Los Angeles Unified School District (LAUSD) during the 2008-2009 school year. The course's effectiveness in broadening participation in computing in Los Angeles has contributed to the acceleration of the number of districts across the country offering this course (see ECS website: http://www.exploringcs.org for more details).

As of the 2017-2018 school year, ECS is offered in 25 states and Puerto Rico, including the 7 largest school districts in the country, as well as many small rural locations. A range of public and private funding sources have been utilized by individual projects to underwrite ECS teacher training across the country. For example, a particular project may provide ECS PD and support local course adoption opportunities funded by NSF grants exclusively, sponsored by other organizations (e.g., Code.org), local school district funding, or braided funding.

ECS has historically used a consistent curriculum across implementation sites. The curriculum has gone through refinement over the years: ECS curriculum version 8 is now available as well as a Spanish version. The structured, two-year professional development program has been a site for design research and an evidence-based model of teacher learning is in place (Ryoo, Goode & Margolis, 2016; Goode, Margolis, & Chapman, 2014). Districts and states who have formally implemented ECS have worked closely with the Into the Loop team in planning and implementing the PD workshops and curricular adoption in local educational agencies. Into the Loop team members provide outreach to projects interested in applying for NSF grants as well as well-prepared facilitators to lead PD sessions. Nonetheless, there have been changes over time in both funding and district-specific policies related to teacher certification and course requirements that have required some shifting on the part of individual projects and priorities.

CSPd Week Summer 2017. The second annual CSPd week, a five-day event held at the Colorado School of Mines, was offered in July 2017. CSPd Week is a residential professional development (PD) program involving parallel sessions of CS PD. ECS partnered with Bootstrap, AP CSP and C4C on the implementation of the second annual CSPdWeek. In addition to funding from NSF, CSPdWeek was sponsored by the Infosys Foundation, the National Center for Women & IT and the Computer Science Teachers Association. Colorado School of the Mines generously hosted the event. Organizers made particular efforts to recruit rural teachers and other teachers who are not involved in large urban CS partnerships to participate.

One hundred twenty-two teachers from 29 states and the District of Columbia attended ECS CSPdWeek Summer 2017 (48 of whom were attending their second summer of ECS PD). CSPdWeek was established to create a multi-faceted CS educator community that provides professional development opportunities to teachers as well as school counselors to help students gain access to education and career opportunities in computing. The event was modeled after residential summer camps, in which participants sign up for a "track" and participate in larger evening events designed to build a learning community and share content and information across tracks.

Key Findings

CS10K Projects

- **50 projects** were funded under the CS10K umbrella since program inception in 2012. Of this number, **45 provided teacher PD** in a manner consistent with EWG collection. Of the 45 teacher PD projects, **44 projects** submitted at least partial datasets over the past three collection cycles.
- **28 projects** offered PD for Computer Science Principles (CSP) and **18 projects** offered PD for Exploring Computer Science (ECS) (two projects provide PD for both ECS and CSP).
- **27 projects** trained new teachers in ECS or CSP for the 2017-18 academic year. Of these, **23** (85%) reported at least partial data for 2017-18.

CS10K Teachers

- 3,255 teachers have participated in the 45 CS10K funded PD programs between 2012-2018.
- Half of teachers are female (50%); and most teachers identified as White (78%) and non-Hispanic (91%).
- 68% of teachers entered PD with at least six years of K-12 teaching experience in any subject; 83% of teachers were new to teaching computer science.

CS10K Students

- CS10K projects currently receiving CS10K funding or "active projects" report reaching 23,708 students during the 2017-18 academic year. This number represents students reached by the total number of teachers for whom student data is reported (14% of teachers in the CS10K count).
- We estimate more than 521,000 students potentially have access to a CS10K teacher in 2017-18. (This is the number of students enrolled in schools with at least one CS10K-trained teacher). This is nearly 4% of the high school student population in the United States.[10]

CS10K Schools

- Over the past three years, CS10K trained teachers have taught in more than 1,653 schools across all 50 states, the District of Columbia, and Puerto Rico.
- In 2017-18, 503 schools added at least one newly trained CS10K CS teacher.

[10] https://nces.ed.gov/fastfacts/display.asp?id=372

Chapter 2: Findings

This chapter presents the key findings are summarized below, followed by the methods and limitations and then the findings in greater detail. Appendix B provides details about the methods and limitations and Appendix C presents a set of data tables for all years of data collection.

CS10K is only one of the PD programs preparing teachers to offer rigorous computer science instruction. There are likely thousands of teachers reaching hundreds of thousands of students to offer a high quality computer science experience, through integrated computational thinking or stand-alone courses and across the K-12 spectrum. This report considers the impact of CS10K on the computing environment.

Methods and Limitations

This represents the fourth year that we have collected data from those CS10K projects providing direct professional development to CS teachers. For the purposes of our collection, this includes PD for CSP, ECS, or other rigorous approaches to teaching CS. we did not attempt to collect any data from the five CS10K projects that were not designed to provide PD (i.e. CS10K projects that primarily focused on CS research or providing resources).

During the 2017-18 data collection year, we encountered additional challenges with data collection due to the changing nature of projects:

- Over half of the projects originally funded in 2012 had sunset and were therefore not in a position to continue providing data.
- Approximately 20% of CS10K projects had secured additional grant funds (typically under a separate award) to continue their work.
    - Some of the projects were finishing their initial funding cycle concurrent with new funds but were reporting only under one project name.
- At least two projects had braided their NSF funds with other available funding. It was therefore impossible to fully disentangle in order to attribute teachers trained to a specific funding source.

The EWG is in the second year of piloting data collection with CSPd Week/Pathways Summer Institute which offers intensive PD for teachers over the summer which is partially supported by CS10K (including using CS10K influenced curriculum and trainers) and other, non-government funding streams. It is likely there are other similar teacher PD efforts in which the NSF investment may not be wholly responsible for the offering and yet highly influential on the outcomes.

Table 2 shows the number of projects funded per year and the number of projects reporting in each of the EWG data collection years. Of note:

- Of the 50 funded projects, 20 focus on ECS and 27 focus on CSP, 3 focus on both. Five are research or resource focused.
- 27 out of 50 projects were expected to submit data; others are either no longer active or are not focused on PD. Several projects are reporting under new funds for training completed as

part of CS10K; others have found additional funding sources to support efforts. Thus, the EWG included the number of projects for which we anticipated data to be submitted.

- 23 Projects submitted data; however, several projects are reporting multiple awards as a single project for a total of 20 unique submissions for the 2017-18 academic year.

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

Award Year	Number of Awards	Projects Reporting to EWG			
		2015	2016	2017	2018
2012	15	12	12	3	3
2013	10	4	6	8	3
2014	12	-	11	7	7
2015	6	-	-	6	5
2016	7	-	-	1	5
Total	50	16	29	25	23*

**Several projects reported data as part of new or partner projects (for example, the 2015 Online Professional Development for Exploring Computer Science is captured as part of CS PD Week.) There are 20 unique project submissions.*

As shown in Table 3, data collection for 2017-18 was largely successful with all but four projects reporting to the EWG the number of teachers trained.

Table 3: Number of CS10K projects reporting data, by collection year (based on unique submissions*)

Collection year	Number of CS10K projects expected to report	Number of CS10K projects that provided data	Teacher level data				Student-level data		School-level data
			Demo-graphics	Disability status	Experience	Implemen-tation	Demo-graphics	Disability status	
2014-15	23	21	21	12	17	11	12	5	18
2015-16	29	26	22	20	18	16	18	11	26
2016-17	26	25	21	18	21	17	15	7	19
2017-18	27	20	20	14	18	14	15	7	17

** Though 23 projects reported data, there were only 20 unique submissions (some projects were funded under multiple grants). Data to the right of this table are based on the 20 unique submissions.*

2014 was the first year the EWG asked for these data. Data from 2011-2014 required historical reflection and, as such they had to be reconstructed by evaluators (and evaluators could not retrospectively incorporate EWG-suggested metrics to collect the appropriate data). Data available were not necessarily parsed by project year, therefore 2011-2014 is presented cumulatively.

Participation in this data collection effort was voluntary. Though most projects submitted data (see Appendix D for the data submission form), not all projects were able to provide data for each element, and it is unclear if the data provided represents 100% of the teachers trained, or a subset of whom

participated in the project-level evaluation. The findings presented in this report reflect the available data, which as shown is incomplete for at least some projects and categories. This is our "best snapshot" of CS10K projects, but may not provide the full picture. Our analysis of the data also assumes all CS10K projects are equally efficacious, as the EWG has no way of evaluating the relative strength of a particular project. Finally, there may be biases in the data based on the size of the projects—with larger projects dominating trends for some categories of trend analyses.

More information about the data collection methods and limitations can be found in Appendix B.

Caveats

When examining the data presented it is important to keep several caveats in mind:

- There are more CSP projects and teachers in the overall data set, but there are roughly an equal number of students represented between CSP and ECS. It is important to note that there are significantly more CSP projects and teachers in the overall dataset, and yet there are more students in the ECS projects that have been accounted for. This may be due to better and more systematic data collection than the CSP projects have the capacity for, smaller classes, or the "newness" of the course.
- ECS has traditionally been concentrated in major urban areas, unlike CSP which is a national offering. ECS is now offered in 25 states and largely concentrated in urban districts based on CS10K funding. Thus, the available student data may be overly representing specific regions and not be representative of all students reached through CS10K.
- It is difficult to track the extent to which trained teachers have moved to different schools or were unable to implement a CS course. The list of schools to which teachers are assigned is cumulative and evaluators report data derived from either the application forms (most common) or through program records, which are updated by the project. Most projects confirm that the schools they provided to the EWG are schools in which the teacher is actively teaching CS. The EWG suspects, however, that over the four years of data collection a significant number of teachers may have moved schools, or are unable to implement a course despite best intentions (e.g., course may be cancelled to low enrollment, competing priorities at a school or personal problems), thus, the list likely has some schools in which a teacher is no longer teaching and/or include schools in which there is a teacher but the teacher is not implementing the course as planned.
- It is unknown how many teachers participate in more than one PD opportunity. It is possible that a subset of teachers are prepared in both ECS and CSP through different PD providers.
- Student data is underreported. At best, students reached are reported for 14% of the teachers trained. Student data are perhaps the trickiest data to collect given IRB/RRB requirements. For the projects that have successfully collected data, it is typically done through district records and/or direct student surveys. Aggregated data provided to EWG that was collected via direct student surveys are limited to those for which there is parental consent/student assent; therefore, the data provided is often for a subset of the students reached. There are several projects (exact number unknown) that collect the student data through teacher surveys (although teachers are asked not to "guess," but to draw student demographics from administrative records).

Findings

This report presents the cumulative impact of the CS10K program where possible. Specifically,

- All results concerning teachers are aggregated across data collection years (cumulative).
- Student data are provided for the 2017-2018 academic year only.
- Data about schools are aggregated across data collection years (cumulative).

Teachers

Q1: How many teachers have been trained in CS10K projects?

The 45 CS10K projects that have reported on their PD efforts since the 2011-12 academic year have provided professional development to 3,255 teachers. As shown in Figure 1, the number of teachers participating in PD increased each year to a high of 918 teachers during the 2016-17 academic year and dropped slightly for the 2017-18 academic year, possibly because there were fewer active projects in 2017-18. Similarly, at least some of this annual growth, particularly between the 2016 and 2017 data collection years may be due to the increase in the number of funded projects, as well as the enhanced capacity of projects to report participation data as part of our collection process.

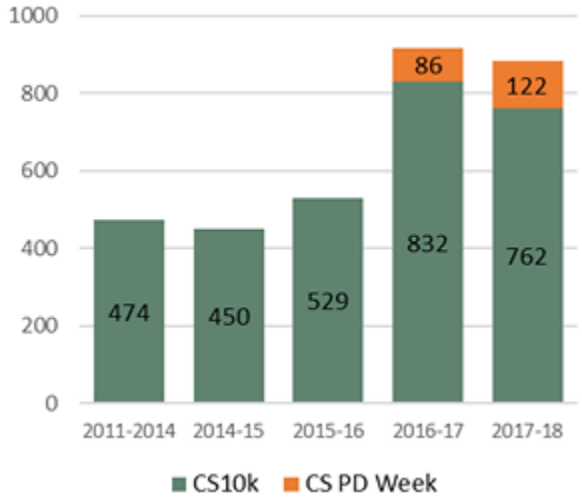
About the data contributing to Q1. *How many teachers have been trained through CS10K?*

Number of projects reporting	45
Number of teachers	3,255
Number of ECS teachers	1,085
Number of CSP teachers	2,170

Number of teachers trained in*:

2011-14	474
2014-15	450
2015-16	529
2016-17	918
2017-18	884

Figure 1: Teacher participation in CS10K over time 2018



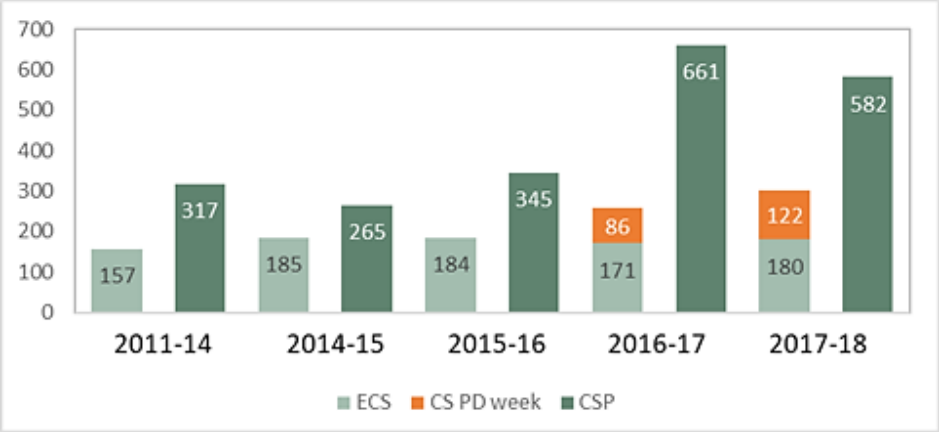
The leveling off of teachers trained between the 2017 and 2018 data collection years is not well understood. Though there are fewer active and reporting projects, the number of teachers trained between 2016-2017 and 2017-2018 is largely consistent. Open questions include:

- Are funded-projects exhausting the supply of local teachers?
- Are other PD providers, available curricular materials, or other supports that are expanding the options for CS PD? Are districts directly funding PD for their teachers?
- Are there teachers being recruited into other PD programs that are not designed for a stand-alone course like ECS or CSP, but integrating CS into existing subjects (supported by the STEM+C program)?
- Are there systemic issues for getting teachers and in and through PD, for example garnering district support for participating in CS PD?

Two-thirds of the teachers receiving CS10K PD were to prepare to teach CSP (2,170). An additional 1,085 teachers were trained in ECS. As shown in Figure 2, the number of teachers receiving training for CSP exceeded the number receiving training for ECS in each of the academic years covered by our collection, and while the number of teachers trained for CSP grew through 2017, the number trained for ECS remained largely constant over the last four years for projects excluding CS PD week. This is not surprising, given that more CS10K projects have been funded to prepare teachers to offer CSP than ECS and that CS PD Week is designed to scale up ECS training.

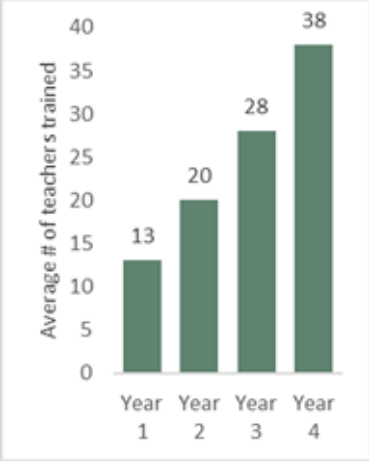
Note that Figures 1 and 2 include a group of 208 ECS teachers who attended Computer Science Professional Development Week (CSPd Week; highlighted in orange and counted as part of ECS) through the Into the Loop and Online Professional Development for Exploring Computer Science projects. Though teachers participated in PD for ECS or CSP, our data only includes CSPD Week participants who were prepared for ECS.

Figure 2: Teachers trained by course



The average number of teachers trained per project seems to increase each project year (Figure 3).

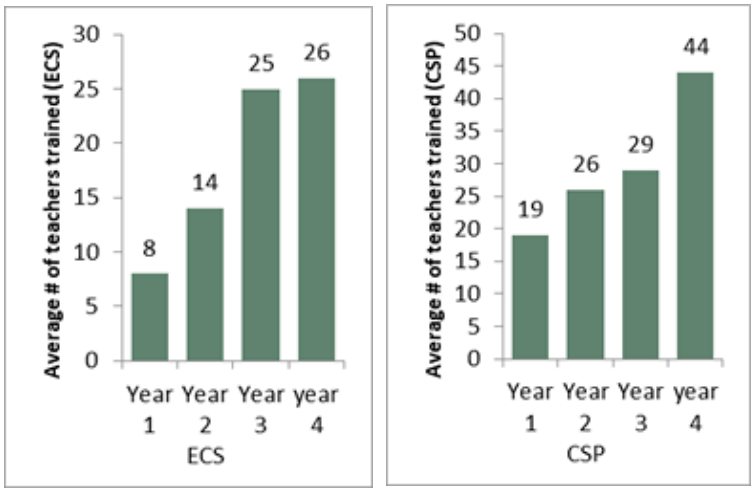
Figure 3. Average number of new CS10k teachers by project year



There is a difference in the numbers of teachers trained per project year between ECS and CSP projects, as shown in Figure 4. ECS and CSP both start with a small set of teachers before scaling up. In ECS projects, this number holds steady at 25-26 teachers in years 3 and 4, while CSP projects peak dramatically in year 4.

Figure 4: Average number of teachers trained per project year by course

By design, ECS training sessions target 24-26 teachers per session, which may attribute to this pattern, however it is impossible to know for how many training sessions each project are reporting data. ECS projects also range from reporting on 2 teachers to 122 teachers. It is also possible that with the formal introduction of the CSP AP exam demand for training increased in recent years.



Q2: How diverse (race / ethnicity / gender / disability status / teaching experience) are the teachers who have been trained through the CS10K program?

This section reports teacher demographics as an aggregate across all project years. To see the demographics broken down by project year please see Appendix B.

About the data contributing to Q2. How diverse are the teachers who have been trained through the CS10K program?

	Projects reporting thru 2018 (N=45)	Teachers counted thru 2018 (N=3,255)
Demographics		
Gender	43	2,816
Race	42	*
Ethnicity	40	2,563
Disability	36	2,063
Teaching experience		
Any K-12	40	2,337
Any CS	38	2,134

Gender

Data on gender were provided by 43 projects for 2,816 of the 3,255 (87 percent) teachers who provided data over the four years of our collection. Of this number, teachers were nearly evenly split between identifying as female and as male (See Table 4) with no significant differences when comparing the ECS and CSP teacher samples. Fifty percent of teachers overall identify as female and 49% identify as male, with one percent preferring not to answer. There has been no significant change in the percentage of male versus female teachers reached over time.

Race and Ethnicity

Data on race and ethnicity were provided by 40 and 42 projects respectively for 2,654 (82 percent) and 2,563 (79 percent) respectively of the 3,255 teachers who provided data over the four years of our collection. Across teachers who participated in CS10K projects and from whom data is available,

most identified primarily as white and non-Hispanic or Latino. Nine percent of teachers identified as Black or African American, and 8% identified as Hispanic (See Table 4). However, for ECS, 15% of teachers identified as Black and 71% as White, compared with 8% Black and 81% white for the CSP teacher sample. As Table 4 indicates, ECS projects were, overall, more likely to have teachers who fell into URM categories than did CSP projects, though the difference was not statistically significant.

Table 4: Percent of teachers teaching ECS and CSP reporting being from an underrepresented racial or ethnic group

% URM ECS Projects (N = 17)	% URM CSP Projects (N = 27)	Difference	Significance Value
29%	19%	10%	P = .13

Disability Status

Over the last four years of data collection, projects have increased their capacity to report on disability status. Data were provided by 36 projects for 2,063 of 3,255 teachers (64 percent; cumulative across 4 years of data collection). Of the teachers for whom data on disability status are provided, 92% reported no disability, 5% reported a disability and 3% preferred not to answer this question (See Table 5).

Table 5: Gender, race, ethnic identity (across all years of data collection)

	Number	Percent
Total number of teachers	3,255	n/a
Gender		
Female	1,417	50%
Male	1,370	49%
Prefer not to answer	29	1%
Total	2,816	
Race		
American Indian or Alaska Native	35	1%
Asian	132	5%
Black or African American	274	10%
Native Hawaiian or Other Pacific Islander	8	<1%
White	2,065	78%
Other	67	3%
Prefer not to answer	73	3%
Total	2,654	
Ethnicity		
Hispanic or Latino	179	7%
Not Hispanic or Latino	2,333	91%

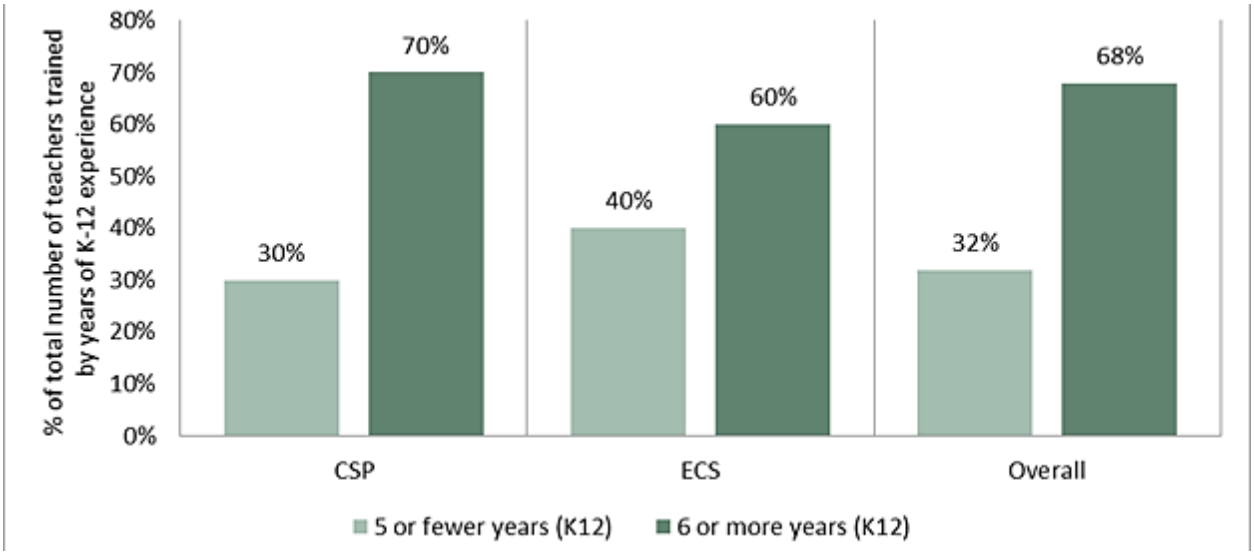
	Number	Percent
Prefer not to answer	51	2%
Total	2,563	
Disability		
Yes	102	5%
No	1,903	92%
Prefer not to answer	58	3%
Total	2,063	

Note: Percentages are based on total number of participants in a given demographic category for which data were available.

Teaching Experience

K-12 Teachers new to CS10K are, overall, an experienced group of K-12 teachers. As Figure 5 shows, for the 40 projects for whom data were reported, 68% of teachers (across all years) have at least 6 years of K-12 teaching experience. CSP teachers are slightly more likely to be veteran K-12 teachers when compared to ECS teachers. The K-12 experience of new-to-CS10K teachers remains fairly consistent over the life of the aggregated projects. This section reports teacher credentialing and experience as an aggregate across all project years. To see the data broken down by project year please see Appendix B.

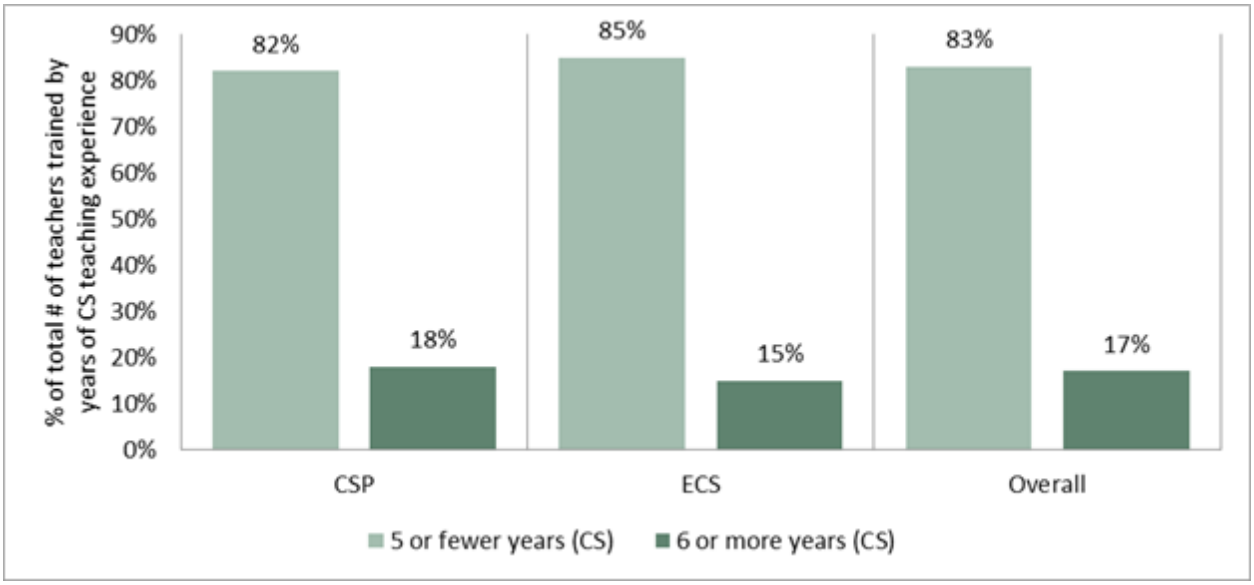
Figure 5: K-12 experience by PD type (across all years of data collection) 2018



Experience Teaching CS Courses

Though teachers are typically experienced in K-12 classroom instruction, they tend to be novice CS teachers. For the 38 projects for whom data were reported, 83% have taught CS for less than six years, if at all (see Figure 6) with no meaningful difference between ECS and CSP teachers.

Figure 6: CS experience by PD type (across all years of data collection) 2018



Q3: In what fields are the CS10K teachers certified and/or teaching?

About the data contributing to Q3: Teaching Fields		
	Credentials held	Primary teaching field[1]
Projects reporting	31	27
Teachers counted*	2,182	1,237

Table 6 shows the fields in which teachers hold a credential. There are 31 projects that have reported data on 2,182 teachers. There has been a trend towards a slight decrease in teachers that specifically hold a math credential but overall the numbers have remained consistent. Interestingly, this shows that CSP teachers are more likely to hold a math credential, and ECS teachers are more likely to hold a science credential.

Table 6: Teaching credentials held

	*Cumulative * N = 2,050*	*ECS N = 463*	*CSP N = 1,587*
Math	26%	20%	28%
Computer Science	15%	11%	17%
Business	15%	15%	14%
Career or tech. ed.	15%	14%	15%
Science	10%	16%	9%
Other credential	11%	14%	10%
English / Social Studies	7%	7%	7%
No credential or certificate	2%	3%	1%

**Teachers could select more than one option*

Data were reported by 27 projects about primary teaching field (subject in which at least 50% of courses are taught) for 1,237 teachers. Table 7 shows the fields in which teachers are primarily teaching. The largest percentages of teachers are both credentialed in and teaching math; however, there is a clear difference between ECS and CSP teachers. ECS teachers are more likely to teach Career and Technical Education courses while CSP teachers are more likely to be dedicated CS teachers or teaching two or more disciplines. It is possible that the distinction is due to a) the increase in CS offerings in high schools generally, b) the requirements for how ECS is "counted" in a school or district, particularly for those seeking to leverage Perkins funding and/or c) ECS being perceived as more accessible to a range of teachers while CSP requires greater content and programming knowledge.

Table 7: Primary teaching field (At least 50% of courses are taught in this subject area)

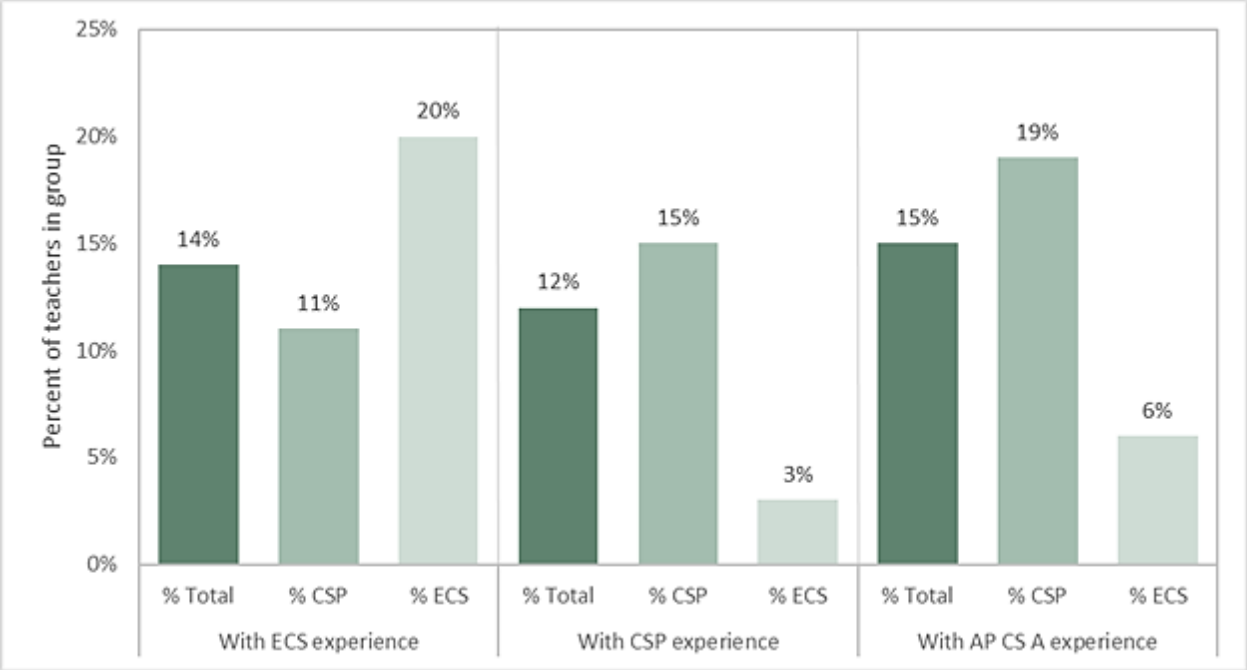
	Cumulative Through 2017-2018		
	Overall N = 1,237	ECS n = 202	CSP n = 1,035
Math	22%	19%	22%
Computer Science	17%	18%	29%
Career or tech. ed.	14%	20%	12%
Science	8%	15%	6%
Teach two or more disciplines	14%	5%	16%
English / Social Studies	4%	5%	4%
Other discipline	3%	8%	2%
Business	7%	8%	6%
Not currently teaching	1%	1%	1%

In 2015-16, the EWG was eager to learn if there is a difference between teachers who are preparing for ECS or CSP. The EWG hypothesized that CSP teachers are more likely to come from a math or CS background (and possibly have a longer history of teaching AP courses) and that ECS projects may have more of a representation of career and technical education teachers than the CSP projects, but data was too limited to draw conclusions. Based on current data, this hypothesis appears to have some merit- both courses are taught by math and CS teachers, though CSP has more CS teachers or teachers who teach in two or more disciplines (perhaps math and CS) and ECS draws more heavily from individuals who primarily teach CTE, and science than CSP does.

In the first year of the EWG common data collection effort, teachers were asked how many years of experience they had teaching a variety of specific CS courses. The question was subsequently pared down to ask if they had any experience teaching three specific CS courses. The cumulative data shows that about 14% of teachers have taught ECS, 11% have taught CSP, and 20% have taught AP CS A (See Figure 7). ECS PD is designed to be offered over two summers, both before and after the first year of

implementation, so it is possible that some projects gathered this information for teachers who are in their second summer of PD after having implemented the course once.

Figure 7: Experience teaching CS courses by total and PD type 2018



In general, CSP are more experienced CS teachers than the ECS teachers are; however, ECS teachers are more experienced in teaching ECS prior to training). Data may have been collected for teachers during their second year of PD. This may be the case for CSPd Week in particular. As has been the case in the past, ECS teachers are more likely than CSP teachers to have prior experience teaching their subject course.

Q4: How many CS10K teachers are teaching with the instructional materials/approaches/ curricula they have been trained to teach?

About the data contributing to Q4: *Implementation of PD?*		
	Projects Reporting 2018	Teachers counted 2018
ECS	12	500
CSP	23	1,537

We asked evaluators to gather information about the extent to which teachers were utilizing the PD materials within the class for which it was intended (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 prior to the first EWG data request in 2015; therefore, the figures reported are for the most recent three academic years only (see Table 8), and only for those projects that returned data to us in the standardized data shell format.

Of the 1,207 teachers trained in projects reporting data about implementation, 58% teach only or primarily using the materials/approaches/curricula from their professional learning experiences

along with supplementary materials, and 8% use none of the materials from their experience. Another 15% report using some material from the PD but relying mostly on other materials.

Table 8: Implementation of PD, across all years of data collection

	Overall *n = 1,207*	ECS *n = 167*	CSP *n = 1,040*
Teach using only materials from training	34%	19%	36%
Teach primarily with materials from training	24%	48%	20%
Teach about half with training materials and half with other	9%	17%	7%
Teach primarily with other materials, use training materials to supplement	16%	5%	17%
Do not use any training materials in teaching course	8%	7%	8%

The EWG did engage in a discussion as to why there might be a difference between ECS and CSP teachers (see Table 9) and the level of implementation reported, though this preliminary conversation is purely speculative. The EWG recognized that while ECS is a curriculum, CSP offers a Curriculum Framework, though some of the CSP projects do offer a curriculum or other resources. For example, one CSP initiative was piloting the course itself, allowing teachers to draw teaching resources from an open array of resources. Another CSP initiative offered a complete curriculum, while a third offered activities, resources, and a pacing guide. Finally, a fourth CSP initiative simply offered pedagogical structures to enhance CSP teaching strategies (but no content materials). As the CSP course and infrastructure mature, more "whole package" curricula (such as Beauty and Joy of Computing[11], and code.org[12]) are being approved and made available. Though there are some differences, this item is one of the few that reveals classroom implementation, a factor that CS education evaluators, researchers, and others evaluating professional learning have struggled to quantify.

[11] https://bjc.berkeley.edu/
[12] https://code.org/educate/csp

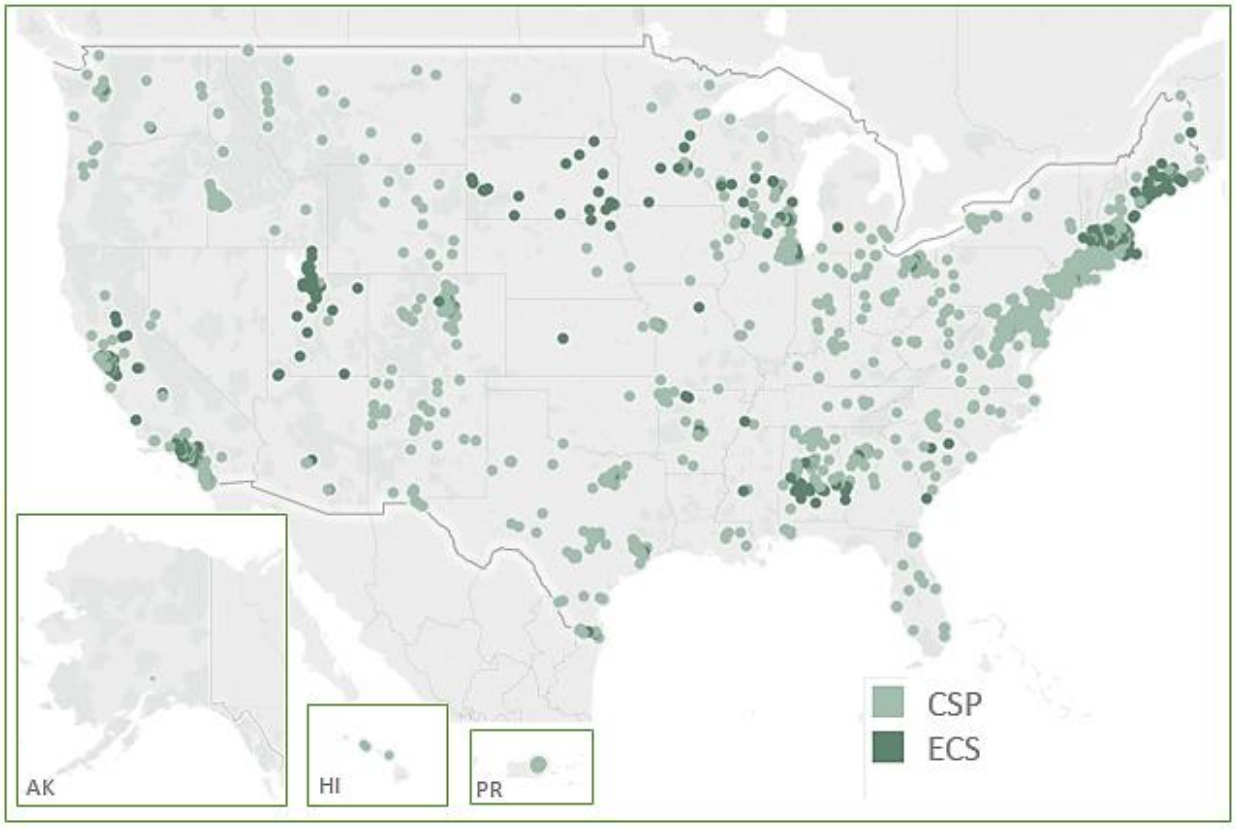
Q5: How many schools have well trained teachers?

As part of the CS10K common data collection effort, the EWG collects a list of all schools in which a teacher is implementing either ECS or CSP. Through data validation interviews, we know that this list may not be completely accurate as some projects collect this information at the time of the teacher application (teachers may subsequently move schools, or a course may be cancelled for a variety of reasons including low enrollment or personal reasons). Thus, the number of schools is our best effort at understanding which students have access to a trained CS10K sponsored teacher.

About the data contributing to Q5: *How many schools have well-trained CS teachers?*	
Projects Reporting	37
Number of Schools	1653
Number of CSP schools	1393
Number of ECS schools	535

Over the life of the CS10K program, at least 1,653 schools have gained at least one trained CS teacher. Figure 8 provides a map of all the locations of CS10K schools.

Figure 8: Location of CS10K schools by PD type



Over the life of the program, all 50 states, the District of Columbia and Puerto Rico have had at least one CS teacher prepared through the CS10K program. Table 9, below, provides the reach of the CS10K program into states and school districts through CS10K and if there is an ECS and/or CSP presence in the state based on CS10K funding.

Table 9: Presence of CS10K Trained Teachers by state

State	Number Projects training in State	Active Projects by Course		Total Number of CS10K Schools	Schools by Course		Total Number of CS10K Districts	District by Course	
		CSP	ECS		CSP	ECS		CSP	ECS
AK	1	1	0	1	1	0	1	1	0
AL	9	7	2	86	62	24	40	40	1
AR	9	7	2	60	40	20	17	10	9
AZ	3	2	1	5	3	2	1	1	0
CA	16	12	5	271	144	127	70	45	36
CO	10	8	2	112	92	20	17	14	7
CT	7	6	1	69	68	1	31	30	1
DC	5	3	2	7	5	2	3	3	1
DE	4	4	0	59	59	0	16	16	0
FL	6	6	0	17	17	0	8	8	0
GA	6	5	1	22	19	3	9	7	3
HI	2	2	0	3	3	0	0	0	0
IA	3	3	0	4	4	0	3	3	0
ID	6	6	0	72	72	0	17	17	0
IL	11	8	3	119	31	88	13	12	2
IN	3	3	0	16	16	0	15	15	0
KS	3	2	1	5	4	1	2	1	1
KY	6	6	0	10	10	0	7	7	0
LA	4	4	0	6	6	0	5	5	0
MA	8	6	2	101	43	58	69	26	54
MD	9	8	1	62	61	1	15	15	0
ME	6	5	1	54	23	31	29	14	17
MI	6	4	2	13	10	3	5	4	2
MN	4	3	1	20	12	8	16	8	8
MO	6	4	2	13	9	4	6	6	0
MS	2	1	1	2	1	1	1	1	0
MT	1	1	0	21	21	0	0	0	0
NC	6	5	1	22	21	1	12	12	1
ND	1	1	0	1	1	0	1	1	0
NE	2	2	0	3	3	0	1	1	0
NH	4	4	0	10	10	0	9	9	0

State	Number Projects training in State	Active Projects by Course		Total Number of CS10K Schools	Schools by Course		Total Number of CS10K Districts	District by Course	
		CSP	ECS		CSP	ECS		CSP	ECS
NJ	9	9	0	51	51	0	33	33	0
NM	5	5	0	40	40	0	27	27	0
NV	1	1	0	2	2	0	2	2	0
NY	10	8	2	90	84	6	40	40	0
OH	7	6	1	56	52	4	34	32	4
OK	1	1	0	0	0	0	0	0	0
OR	3	3	0	5	5	0	2	2	0
PA	7	7	0	28	28	0	16	16	0
PR	1	1	0	0	0	0	0	0	0
RI	4	4	0	6	6	0	6	6	0
SC	7	5	2	11	9	2	8	7	1
SD	3	2	1	29	2	27	6	2	4
TN	4	4	0	6	6	0	4	4	0
TX	12	9	3	120	107	13	51	48	6
UT	4	3	1	62	5	57	31	2	29
VA	6	6	0	34	34	0	12	12	0
WA	7	6	1	17	15	2	11	11	0
WI	9	7	2	71	43	28	33	16	18
WV	3	3	0	13	13	0	2	2	0
WY	2	1	1	15	14	1	2	1	1

Q6: How many students have been reached through the CS10K projects?

About the data contributing to Q6: *How many students have been reached through the CS10K projects*

	All CS10K students	ECS	CSP
Number of students reached in 2017-18	23,708	12,598	11,110
Number of teachers for whom there is student data reported	459	180	279
Percent of teachers for whom there is student data reported	42%*	91%**	31%**

**There were 15 projects reporting student data in 2017-18. These 15 projects have trained a total of 1,099 teachers (197 ECS teachers and 902 CSP teachers). Student data is available for 459 of these teachers (42%).*
*** 91% of ECS teachers trained by the ECS projects that submitted data reported student-level data while 31% of teachers trained by CSP projects had teachers who reported student data.*

Student data are incredibly challenging to share with third-party evaluators. This is due to challenges associated with local IRBs and RRBs, lack of parental consent to share student data, the cost and staffing associated with complying with a special data request, and difficulty accessing historical student data. Because of these obstacles, only 15 of the 45 CS10K projects that reported on their PD efforts provided any student data. As such, student data for students in ECS and CSP courses with a CS10K teacher in 2017-18 are only available for 459 teachers—representing 42 percent of the 1,099 teachers in these 15 projects, and 14 percent of all 3,255 teachers that participated in the CS10K program since the 2011-12 school year. These 15 CS10K projects provided data on 23,708 students during the 2017-2018 academic year. Of this number, 12,598 (53%) students participated in ECS and 11,110 (47%) were in CSP. It should also be noted that thousands of additional students were likely enrolled in in CSP or ECS courses taught by teachers trained via PD funded via non-NSF initiatives.

Student data are incredibly challenging to collect. The 15 CS10K projects that provided student data provided data on 23,708 students during the 2017-2018 academic year. Of this number, 12,598 (53%) students participated in ECS and 11,110 (47%) were in CSP.

As stated in the challenges section, some projects struggled to provide data on the number of students who participated in a CS course in a given school year. This was due to challenges associated with local IRBs and RRBs, lack of parental consent to share data, and difficulty obtaining historical data. It should also be noted that thousands of additional students were likely enrolled in in CSP or ECS courses taught by teachers trained via PD funded via non-NSF initiatives.

Q7: How diverse (Race / Ethnicity / Gender / Disability status) are the students who have been reached through the CS10K program?

About the data contributing to Q7: *How diverse are the students who have been reached by the CS10k program?*	
Total students (2017-18)	23,708
Total % for whom gender is reported	93%
Total % for whom ethnicity is reported	61%
Total % for whom disability is reported	57%
Total % for whom race is reported	51%

Data on the demographic makeup of the 2017-18 student population reached by CS10K teachers was provided for five out of eight ECS projects and ten out of twelve of the CSP projects (See Table 10), yet nearly half of the data was provided by one ECS project. Related, there are more students reported by ECS projects though there are more teachers reported having been prepared to offer CSP. This may be due to better and more systematic data collection than the CSP projects have the capacity for, or the concentration of ECS projects in a smaller number of school districts. Regardless, the diversity of students should be interpreted with extreme caution. This section reports student demographics for only 2017-18 project year. To see the data broken down by project year please see Appendix B.

The data suggest that the CS10K is working to increase participation in computing, as at least 23,708 students were reached in 2017-18 through teachers trained by CS10K. Digging into the demographics of these students suggests there is still work to be done to reach both gender and racial/ethnic parity between projects and across courses. For example, the EWG notes that:

- 43% all student data comes from one ECS project located in an urban area (Chicago)
- ECS may be a required course, while CSP is an optional course
- CSP is offered only in schools with an AP program

Table 10: Number and characteristics of students reached in 2017-18 by teachers trained through CS10K

	CS10K		ECS		CSP	
Total number of students reached	23,708		12,598		11,110	
	N	%	N	%	N	%
Gender						
Female	8,780	39%	5,548	44%	3,232	33%
Male	13,163	59%	6,842	54%	6,321	64%
Prefer not to answer	10	<1%	-	-	10	<1%
Data not available[13]	453	2%	208	2%	245	2%
Total	22,406		12,598		9,808	

[13] "Data not available" is a distinct category from missing data. This category represents project-reported data not available, possibly due to not collecting certain demographics at the student level, versus the field being left blank in the submitted data shell. Missing data accounts for the remainder of the total to equal 23,708.

	CS10K		ECS		CSP	
Race						
American Indian or Alaska Native	132	1%	69	1%	63	1%
Asian	1,227	6%	617	5%	610	8%
Black or African American	3,870	20%	3,092	27%	778	10%
Native Hawaiian or Other Pacific Islander	38	<1%	18	<1%	20	<1%
White	6,305	32%	1,586	14%	4,719	58%
More than one race selected	390	2%	161	1%	229	3%
Prefer not to answer	87	<1%	-	-	87	1%
Data not available	7,502	38%	5,923	52%	1,579	20%
Total	19,551		11,466		8,085	
Ethnicity						
Hispanic or Latino	6,488	41%	5,319	46%	1,169	27%
Not Hispanic or Latino	7,943	50%	5,543	48%	2,400	55%
Prefer not to answer	54	<1%	-	-	54	1%
Data not available	1,401	9%	621	5%	780	18%
Total	15,886		11,483		4,403	
Disability						
Yes	1,683	10%	1,538	13%	145	3%
No	11,770	73%	8,743	76%	3,027	66%
Prefer not to answer	15	<1%	-	-	15	<1%
Data not available	2,634	16%	1,202	10%	1,432	31%
Total	16,102		11,483		4,619	

Anecdotally, among the EWG team members that evaluate CSP projects, we have seen that teachers report that their CSP classrooms mimic the racial/ethnic breakdown of the school but are working towards closing the gender gap. There are early indications that the gender gap gets smaller over time as teachers are able to recruit, and word of mouth spreads.

The EWG has heard informally from PIs and members of the CS education community that teachers are hoping that their schools develop an ECS → CSP pipeline. With more students being exposed to computing through ECS, teachers hope that there will be students that move onto CSP that may not have otherwise pursued a computing course. Some teachers wonder if there will be CSP students that then also take the AP-CS-A exam; however, the courses are seen as quite distinct as CSP focuses on computing more broadly and CS-A focuses on programming. Utilizing statewide data systems could shed light on any CS pathways that are being established.

Q8: How many and how diverse (race/ethnicity/gender/disability status) are the high school students potentially reached through the CS10K program?

We estimate more than 521,000 students potentially have access to a CS10K teacher in 2017-18. (This is the number of students enrolled in schools with at least one CS10K-trained teacher). This is nearly 4% of the high school student population in the United States, compared to nearly 6% of the

high school student population potentially reached in 2016-2017). This shift is likely due to underreporting.

The EWG collected data on a nearly even number of CSP students and ECS students reached through CS10K. When we look at the data as provided by projects in 2017-18 against the data for the schools in which we know CS10K teachers are offering CSP and ECS we see that the CSP teachers are in schools that typically reflect the national population, with the exception of fewer students receiving free or reduced lunch. Schools offering ECS have a higher percentage of Hispanic or Latinx students and Black or African American students than the national norms. CS10K schools have a lower rate of free and reduced price lunch than the national average (see Table 11).

Table 11: Students with potential access to a well-trained CS10K teacher, 2016-2017

	CS10K student participation [14]	CS10K School population	ECS schools: total population	CSP schools: total population	National: total population[15]
Number of students	23,708	521,311	126,462	394,282	14,949,714
Gender					
Female	40%	49%	50%	49%	46%
Male	60%	51%	50%	51%	48%
Race					
American Indian or Alaska	1%	1%	1%	1%	1%
Asian or Asian Pacific Islander	10%	7%	8%	6%	5%
Black or African American	32%	17%	26%	14%	15%
Hawaiian or native pacific	<1%	<1%	<1%	<1%	<1%
White	53%	46%	28%	52%	51%
Two or more races	3%	3%	2%	3%	3%
Ethnicity*					
Hispanic or Latino	45%	26%	36%	23%	25%
% Free/Reduced lunch**					
% Free/Reduced lunch	N/A	43%	59%	38%	50%

**Note, NCES does not separate ethnicity and race.*

Q9: How many and how diverse are the students who take the AP CSP exam?

Only 4 projects reported this data, and most of it came from one project which was unable to provide demographic data. Ideally, projects will increase their capacity to track how many students with CS10K trained teachers have taken the AP CSP exam, which will allow us to report this data and

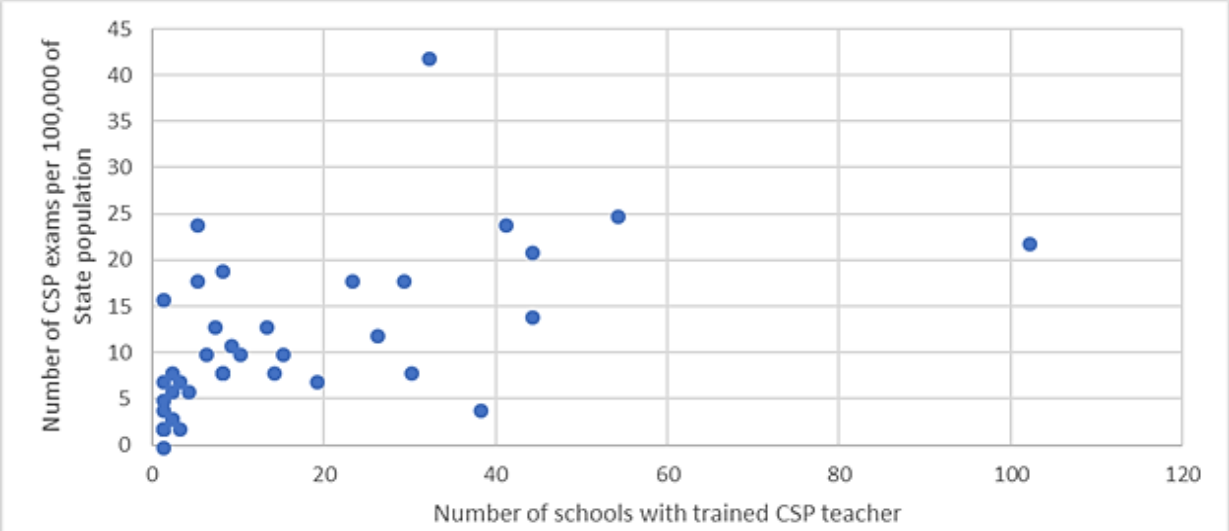
[14] Percentages only reflect students for whom we received data (excludes those who said "no answer," "data not available" or whose responses were omitted)

[15] Note that these numbers reflect the overall population of public school students in grades 9-12, with the exception of the free/reduced lunch percentage, which reflects public school students across all grades. The numbers are slightly different likely due to inconsistent reporting to NCES for race and gender.

compare to data released by the College Board, with a particular eye towards demographic parity of exam takers.

Data collected allows us to explore the extent to which the number of trained CSP teachers relates to AP CSP exams taken. To answer, the EWG attempted to correlate the number of schools offering CSP (based on having a school in our database associated with a CSP teacher) with the number of CSP exams taken per 100,000 population.[16] Among the 41 states for which we had a record of a trained CSP teachers, there was a moderate and statistically significant correlation between the number of schools with a trained CSP teacher and the number of CSP exams taken per 100,000 population (see Figure 9). Note that the data used in this analysis is based on the 2016-17 academic year.

Figure 9: Correlation between schools with a trained CSP teacher and number of CSP exam takers.



Note: Correlation coefficient = .56, significant at p < .01

[16] Based on data collected by Georgia Tech's Institute for Computing education (http://home.cc.gatech.edu/ice-gt/597)

Chapter 3: State Data

In a 2016 NSF Dear Colleague letter, Director France A. Córdova recognized that:

> "The U.S. science and engineering workforce can thrive if women, blacks, Hispanics, and people with disabilities are represented in percentages comparable to their representation in the U.S. population. According to the National Center for Science and Engineering Statistics, we have a long way to go to reach that goal."

Recognizing progress towards these goals requires the collection of detailed data that allows for the rigorous analysis of the characteristics of individuals who are participating in and benefiting from initiatives designed to broaden participation in STEM fields. To address this need, The Committee on Equal Opportunities in Science and Engineering led a 2017 NSF workshop of evaluators, educators and administrators ("Workshop on Assessing Performance and Developing an Accountability System for Broadening Participation") produced a report called Better STEM Outcomes: Developing an Accountability System for Broadening Participation. This document offers a framework of ten principles for developing systems for accountability that includes:

- Conduct a self-study that takes stock of your organization's current broadening participation portfolio and climate (baseline).
- Construct a timeline (near- and long-term) for achieving outcomes articulated by your theory of change, consistent with the institutional mission and strategic plan.
- Identify data and measures that are required- either extant or to be created- to gauge progress organization-wide towards your outcomes.
- Engage stakeholders to define a common agenda and recruit partners to work toward agreed-upon outcomes, disaggregated by demographic, educational and career stages as much as possible.
- Incorporate what has been learned from ongoing longitudinal assessment of your organization's broadening participation programs.
- Appraise the performance of your organization in taking steps toward increasing accountability and institutionalizing democratized science and engineering system
- Be ready to begin again, as accountability for broadening participation is a recursive, iterative and ongoing process.

State data systems that embrace these principles will be well poised to use these systems to reach their broadening participation goals. Because having meaningful data is an essential part of the ECEP strategy for state teams to define, set, and measure goals for broadening participation, the ECEP community identified the 2018 Annual meeting topic to be: How to set measurable goals to broaden participation in computing/measurement of broadening participation. This topic emerged from the planning committee, on which 6 states were represented. It gained more urgency over the last year when it was identified as a priority on the ECEP State Survey (administered by SageFox). Further, the need for better outcome data was identified as a need by the EDC/Westat evaluation of the Broadening Participation in Computing-Alliances program, which found that many of the Alliances were not obtaining data on the characteristics of participants reached through their efforts. Planning for and tracking broadening participation is also becoming a more pressing concern across NSF.

State data is the most feasible way to detect change after the CS10K projects are no longer funded as there are no resources for following teachers and students. Therefore, collecting data through state education systems may be one viable (and more sustainable) alternative to the CS10k model for tracking participation in computing education and ensuring equitable access and participation, but

the data must be collected in a systematic, replicable fashion with common variables across states that can demonstrate change over time.

The Challenge

K-12 districts and schools are often asked to provide data about participating teachers and students while a professional development initiative is underway. However, once funding for the initiative ends, districts and schools may lack the resources to continue providing the data required to assess the presence of longer-term outcomes. As such, state data may be the only way to detect change after the projects are no longer funded as there are no resources for following teachers and students.

Tracking participation in computing, setting goals for reaching gender, racial, ethnic and socioeconomic parity, and tracking progress towards parity relies on state-level data systems. The data collection effort through CS10K has provided valuable information, but the community-driven nature of this approach has presented challenges, namely: the reliance on voluntary participation of evaluators, data are not necessarily reliable or complete, and the data are aggregated to the project level, precluding valuable analysis of the impact of PD for teachers and students.

The 2016-2017 data collection year revealed greater challenges with data collection than in previous years due to the changing nature of projects:

- Some earlier cohort projects have sunset and are therefore not in a position to continue providing data.
- Some projects have secured additional grant funds (typically under a separate award) to continue their work. Some of these projects are finishing their initial funding cycle concurrent with new funds, but are reporting only under one project name.
- Some projects have braided their NSF funds with other available funding and may be impossible to fully disentangle in order to attribute teachers trained to a specific funding source (for example Pathways or CSPd Week now called Pathfinders which also used private dollars and which is included in our dataset).

Through a survey and a follow-up EWG-ECEP partnership around the annual meeting, state teams were asked to engaged in deep thinking about state level data to help shape including:

- Procedures for accessing state data systems
- Defining broadening participation goals with measurable objectives
- Identifying metrics for monitoring progress against objectives
- Identifying capacity issues with turning data into knowledge

This report discusses findings and implications from this exercise in detail. Conclusions and recommendations on next steps are then discussed.

Findings

Collecting statewide data is challenging, as many state teams are learning. Teams are sometimes using informal channels to make data requests, are gathering data in a piecemeal manner, or are finding that the systems available are highly controlled and inaccessible.

One of the fundamental issues for many states in accessing data is to define what courses "count" as meaningful CS experiences. For states that have standards in place, it is easier to identify which courses meet standards associated with the chosen definition of CS. For other states, inconsistent use of course codes and names make this much more challenging. In addition, even when courses are "on the books" that appear to teach computer science, the content is not always as described.

Preliminary interviews with state ECEP teams helped us to better understand the challenges states are facing in collecting data. Three themes emerged from these interviews: access to the data; data availability; and analysis and publication.

Access. All state ECEP teams interviewed collected some K-12 student data. Overall there were a number of characteristics of data collection systems that included:

- Closed (difficult for external researchers to obtain data)
- Small (a small system with data requests going through one or two central people)
- Large and centralized (a large data collection system, many districts, many variables based on clear, statewide definitions) maintained at the state level
- Large and decentralized (a network of data collection systems, with district-level control over definitions of CS courses and other variables)

Procedurally, all state-level student data systems had a web page, with five of the six states including a data request form and many of those had written instructions or procedures. One state had a very simple online form (about 10 fields to complete), and another state simply offered the names of persons to contact with a written request. The other four states have more extensive data request procedures. Most states had some student-level data or data reports freely available online. Three of the interviewed states have a centralized state-level data system with one point of access. In two states, student-level data are maintained at the district level, and with one of these states, district-level data are not fully available on the state-level system (e.g., teachers and schools are not linked). A final state offers very limited access to student-level data.

In terms of available data:

- Most states have a centralized list of CS courses; one state allows each district to define what a CS course is, so no centralized course codes exist
- Teachers and schools for each HS course (including CS courses) was noted as available by five of the interviewed state data representatives; five states noted that teacher demographics were available (and three of these carry information about credentialing)
- CS course enrollment information was noted as available in four states; two of these states were positive that demographics were available for enrollment, four states were unsure or unclear about that availability
- Available student outcomes, mentioned in five of six interviews, included
    - Course completion information
    - Grades
    - Test scores

Some issues and key considerations with accessing student data were identified through the EWG discussions of the interview findings:

- Privacy issues with student data are extensive. Finding a state data representative to help navigate the systems may facilitate access by developing an appropriate data request.
- For longitudinal research, looking at course-level outcomes can be difficult, given changes in course content, course codes (e.g., CIP codes, NCES subject areas), and course names.
- Accessing student data occurs with state-level systems that seek to protect teachers, districts, and state reputations.
- Fears of comparisons and accountability abound.
- Student data collection and research access to databases are resource intensive. Therefore, the process of obtaining state-level access necessarily involves proving the worth and value of the research endeavor.

Data Analysis and Publication

In terms of data analysis, most states offer data via an Excel spreadsheet (one state uses Google forms). How "clean" the data are may vary from system to system. The level of detail will vary, depending on the state system (what is potentially available) and the details of the request. For example, two states will not report any student level information for a student group less than 6 in number. Ideal times for making data requests vary from state to state some suggested the fall, while others suggested spring or summer. Finally, in terms of publishing findings based on state-level data, two state representatives stated that the data that is shared is already screened for the expectation that it will be shared elsewhere, and four representatives stated they were not sure, or did not explicitly address the question.

Defining broadening participation goals

As articulated in the Dear Colleague planning for and measuring progress in broadening participation is imperative. Even within the ECEP teams, there was variation in reports of how difficult it would be to *define broadening participation* which is a critical first step in shaping what data to request. For some states it may be that stakeholders are not in agreement, for others it could be that they default to CS for All students, conflating access (i.e. a course exists) with equity (i.e. the participation in the course reflects the student population). Despite the challenges, all states report it is valuable for defining broadening participation in their states, with all but two saying it would be highly valuable.

Not only is it challenging to define broadening participation, but it is even harder to define broadening participation goals, even when assessed as an independent task from defining BPC more generally. Yet again, states overwhelmingly report the collection of BPC data as valuable work.

Preparing to collect statewide data

Collecting data from state systems starts with defining a request, which fundamentally requires identifying the data elements to be collected. One of the greatest challenges across the states is *determining what courses to include in the data request.* Defining relevant computer science courses in the absence of consistent use of course titles and codes and/or the absence of standards is a highly intensive process which may require calling each school or district. This can be further challenging when principals or other administrators don't have an accurate understanding of what computer science is or which classrooms are implementing content that align with a specific definition. Identifying computing courses is further complicated by projects that seek to integrate CS and/or computational thinking practices into existing core curriculum courses like math and science.

ECEP state team members commented that:

> *[this data] can easily be gathered and would be valuable to the team if there is consensus on what "counts" as CS.*
>
> *How courses count has not been clearly defined by the [State] Department of Education*

Many states noted that although it would be time-intensive to determine which courses would count as CS at the high school level, it would be nearly impossible at the K-8 level where any computer science offered is often integrated into other subjects, or may be part of a technology unit. One state noted that *"All answers are given for high school classes only. K-8 data is difficult to collect."* As such, for the purpose of this report, we're focusing primarily on stand-alone high school courses such as ECS, CSP or other similar offerings.

When reflecting on how easy it would be to access statewide data, most states had variation in regards to the ease of collecting data about individuals, about schools, about districts and about the state. Often, these challenges were a result of privacy concerns, making it easier to collect data from larger systems (i.e. district and school) though some states also had concerns about the quality of the data. For example:

> *There are elements that FERPA requires specific regulation; therefore, our data team has been reluctant to pull some data, and the state agency does not report individual data on teachers due to state law. These data are available at the state agency, and we are trying to come up with a way to show program results using aggregate data sets*

> *The [State] Department of Education keeps great records of the demographic enrollment in courses, however, it's unreliable due to who fills out this information at the school level. In some schools, it's the guidance counselors, while in others it's the secretaries. There is no training as to which courses count. Those responsible have little incentive to accurately report information.*

> *Data collection for state courses should be easy, but may be a year old*

> *Multiple layers of data from [State] Department of Education. Can be time-consuming. Teacher data is more difficult. Can't get student grades. AP participation and pass rates come from College Board. We'd like to follow students year to year.*

It becomes more challenging to collect more specific data about individual teachers and students participating in CS, with several state teams noting they don't have the capacity and resources to do so. The same systemic and privacy issues apply, plus the additional challenge of the capacity of a person at the state or on the ECEP team to make a request and subsequently analyze the data so it is meaningful for decision making. For example, states report being stymied by:

> *Utilizing Subject and Personnel Codes (and other data sets contained within the [data] Portal) the ability to gather data pertaining to both the core and sub questions can be done with ease; also of value. While most of this data can be retrieved with ease, it will still take time to compile and query specific data sets; some data requests may also need to be submitted before an internal Data Governance Committee depending on purpose/use*

> *This information is simple to gather on a division by division basis, and [ECEP Team] a tracks a lot of this information as it conducts teacher trainings. The state DOE is not set up to collect this information so state aggregate data is difficult to generate*

> *The AP data are available and can be procured in the aggregate fashion for students. The AP score percentages would require a couple of data pulls and some sheet mergers and calculations. This would not be too difficult but would require more legwork. This also goes for AP pass rates. Reports on students' subsequent course enrollments would take some effort to organize; however, we do have pathway completion reports at an aggregate level.*

Data requests and lessons learned

At least seven of the 17 ECEP states have made data requests about participation in CS courses and have data in hand. The experience and type of data requested and received varied widely (some states requested data prior to the ease/value exercise). Several lessons were learned through this effort:

- Most states relied on an individual to champion the request. In most cases this meant that they had an identified representative from the department of education who endorsed the request, or were themselves in a position to pull the data directly. It is worth noting that all of the states that went through formal channels were met with long lag-times or data that was incomplete or out of date.
    - For many of the ECEP states that have yet to work with data, having a person representing the state's Department of Education on their team allowed for conversation about what was available, and how they would use the data. These pragmatic conversations allowed the teams to think through their BPC goals and measurement in a concrete manner.
- How centralized the state's approach to course content is determined how labor-intensive defining the request was. For a few states, ECEP members literally went school by school to verify that the content in the course listed actually met a rigorous definition of Computer Science. In one state, the team queried the state data system with over 20 key words (e.g. computer, technology, media, C++, database, animation, design etc) at which point each course (at each school) identified was manually reviewed for inclusion.
- Making a request is only part of the process; many states struggled with the capacity to interpret the data received. The more specific the request, the easier the data will be to analyze. In some cases, the labor required to clean and validate the data and make meaningful inferences was often a barrier.
- It takes time to make a data request and receive data. Often it is iterative, in which a team will request a bit of data to work with. Once they have become familiar and developed a trusting partnership with the Department of Education, a more comprehensive data request may be defined.
- The capacity to make a request and use the data is a challenge for many states. For many of the states, ECEP awarded mini-grants to support this work which would have otherwise not been done

Finally, several states have begun to work with state-level data starting with what is "easy." This strategy is useful for states that are seeking to build relationships with those in the Department of Education responsible for data warehousing, begin to build additional capacity within their team and with their partners and facilitates discussions for what is meaningful. Sage advice is to begin somewhere, even if it is not as comprehensive as desired.

Discussion

Using state-level data to help define broadening participation goals for computing education and for measuring progress is universally valuable for ECEP members. State data allows states to:

- Understand the current landscape
- Identify where there are gaps
- Create strategies for broadening participation
- Track progress over time.
- Engage stakeholders and build support

It is, however, difficult to prepare a request and to process data into meaningful knowledge for a variety of reasons including:

- Defining computer science and identifying courses
    - Difficulty defining Computer Science may be challenging for some states. When there are not clear standards associated with a CS curriculum, a specific definition is needed; one that is understood and shared by administrators, teachers, and other stakeholders

    - Identifying relevant course codes for inclusion in the data request and ensuring that these course codes accurately reflect what is happening in the classroom
    - Collecting historical data may be problematic due to changes in course names and numbers and may require additional course lookup.
- Difficulty engaging stakeholders to create a shared goal for broadening participation
- Limitations with the data systems themselves
    - Each state's system is unique and sometimes it requires accessing multiple systems to gather all relevant data
    - Courses may not be coded in a universal manner within a state
    - Courses that integrate CS rather than serve as a stand-alone CS course may or may not be identifiable
    - K-8 classes may be defined and tracked in a manner that precludes identifying CS systematically.
- Privacy concerns
    - May impede who can access the data and for what purposes
    - May have a threshold for the number of teachers and students at which they are able to provide information. This could potentially preclude accessing data from areas with small or emerging CS participation.
- Capacity challenges
    - Making the request requires a significant amount of time (see above challenges)
        - A formal request may not actually be the most effective means of requesting data. Political connections (that not all teams have) are important in the process.
    - Processing the data can be time consuming particularly if the data has errors, comes from multiple sources, or needs sophisticated analysis
    - Sharing data in a meaningful manner with specific stakeholder groups requires production of multiple reports.

Next steps

Using state data is a challenging but worthwhile endeavor. Several states are moving towards legislation supporting CS education which may mitigate some of the challenges by creating a well-defined definition of what constitutes CS, boosting the need for better tracking of CS courses and participation and/or measuring progress against a legislative directive. This could further the urgency in developing the capacity to use state-level data.

ECEP has organized its next phase around data systems and use. The support (financial and technical) that ECEP has offered teams has been critical for making data requests and processing and using the findings. Focusing on the utilization of data for landscape reports, research and advocacy will be paramount. One key element that should be emphasized in landscape reports is the need to explore CS course descriptions and classifications as this work has likely never been done before in most states and can yield valuable insights into how CS is being taught and what it looks like in practice.

The RPPforCS project is replicating and expanding the work of the EWG. RPPforCS brings together all NSF-Funded CS For All: Research Practice Partnerships (RPP) to collect participation measures through a similar process as the CS10K common data collection effort (many of the RPPs are teacher PD focused) and also works with teams to help build capacity for using state data as part of their research efforts.

Chapter 4: Lessons Learned and Recommendations

Four cycles of data collection has offered a series of lessons learned about cross-project, common data collection. Specifically, this project has

- Demonstrated a process for collecting data from projects
- Shown the value of a collaborative and iterative process
- Created a tool useful for future data collection efforts
- Measured spread over time for CS10K projects and gained an understanding of issues of scaling and sustainability
- Identified the potential and challenges associated with state data as a means of tracking longitudinal change.

Ultimately, the collection of program-wide data is necessary and good practice**.** CS10K was designed to increase the number of schools with well-trained CS teachers. Without cross-project data collection it would be impossible to determine if the program is worth the investment. The EWG also uncovered a series of challenges and lessons learned about doing this type of data collection. Though the EWG faced significant challenges with data collection and interpretation, they could all be overcome through an agency commitment to a program-wide effort.

Lessons Learned

The four-year CS10K effort resulted in a number of lessons learned that may be of interest to other groups seeking to conduct cross-project, program-wide, or state-level data collection. In particular, we learned about:

- The importance of engaging the community of evaluators
- Implications of collecting K-12 student-level data'
- Interpreting data aggregated by projects for cross-project analysis
- Limitations of cross project approach, particularly those that are grass-roots efforts.
- Collecting data from state agencies (i.e. state departments of education) may offer a more sustainable data collection approach for understanding the impact of CS teacher PD on student participation

Engaging the community of evaluators is critical

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 deemed 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." 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.

Accessing K-12 student-level data can be logistically challenging and resource intensive

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 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 PD 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.

Interpreting data collected across projects is challenged by each project's interpretation of the request and data collection methods

- Some projects don't fit neatly into the data shell structure due to the way they implement their professional development (e.g. CS CAVE, IDOcode) which may have a different cycle of implementation than the academic year and may not have the same relationship with teachers who are trained.
- Projects used different methods to obtain information about teacher demographics. Although evaluators have been able to consistently provide data about teachers' gender, race and ethnicity over the three data collection cycles, they used different methods to obtain this

information. Typically, projects used teacher surveys to capture this information, but occasionally used application or program records. This may affect interpretation of the data reported.

- Projects used different timeframes to report teacher experience. Most projects asked for the participants' teaching experience for both K-12 and CS courses. The data provided to the EWG seem to be evenly split between projects that collect these data on applications or at the start of PD projects and projects that collect these data at the end of the implementation year. This means that the data for some projects may be "off" by a year (for example, a teacher may have 10 years of K-12 experience at the point of entry during Spring 2015 but the data collected by the EWG would ask for their experience *including* the implementation year, which would be 11 years by Spring of 2016).
- Projects used different methods to track teacher implementation. In addition, evaluators were asked if they collected implementation data beyond the first training year. About half of the projects interviewed report that they do track teacher implementation for the duration of the grant. Several reported that it is challenging to do so when the teachers have no incentive to participate as their formal obligation to the program is over (typically after one year). The data collection spreadsheet currently asks evaluators to report on previously trained teachers, but for some projects this is not possible or there are large amounts of missing data.
- Projects use different methods to collect information about teacher implementation. During data validation interviews in Year 2, the EWG learned that the two most common methods were teacher surveys administered at the end of the year or at the end of each unit. In the former, the question used was typically that which was recommended by the EWG. In the latter, the evaluator may have retrofitted data from observation, teacher report (e.g interviews or surveys not using the EWG recommended questions) or other records to fit the EWG common data spreadsheet. This leaves room for evaluator interpretation that may be inconsistent across projects. It is unknown to what extent this remains true for Years 3 and 4.
- Projects made different interpretations of the requested respondent pool for teacher implementation. During data validation interviews in Year 2, the EWG asked projects regarding whom they administered the teacher implementation questions. The EWG intended data only to be provided for teachers implementing. In most cases the projects that were contacted reported an accurate interpretation; however, in at least one case the project had included teachers who are not implementing and reported they were using "none" of the material. We also learned that some evaluators used the data they collected independently through interviews and/or unit surveys to retrofit the data to the EWG question format. Given the challenges with interpretation and the low response rate, the response to this question may not be truly representative. It is unknown to what extent this remains true for Years 3 and 4.
- Projects struggle to collect student participation data. Student data are perhaps the trickiest data to collect given IRB/RRB requirements. For the projects that have successfully collected data, it is typically done through district records and/or direct student surveys. Data provided to EWG via direct student surveys are limited to those for which there is parental consent/student assent; therefore, the data provided is often for a subset of the students reached. There are several projects (exact number unknown) that collect the student data through teacher surveys (although teachers are asked not to "guess", but to draw student demographics from administrative records). The data are also presented in aggregate form, rather than by school, by student or by teacher which precludes the EWG from conducting any outcome analysis (for example, it would be interesting to know if teachers are successful at broadening participation in their courses).

There are limitations with large-scale data collection efforts

- There are projects that overlap between funding sources. For some projects, project PIs view participants as their "own" whether funded by Code.org, NSF or hybrid funding. Other projects are much more distinct, funding streams/ PD activities work independently of one another. This may lead to double counting teachers or under-counting the overall reach of a program regardless of the funding stream. Similarly, as NSF expands the programs supporting teacher PD in CS, it is more likely each year that the number of teachers trained to offer high quality CS instruction is being undercounted in this report.
- Collecting data from projects at the aggregate level (teachers, schools, and students) limits what the EWG is able to answer. If instead of aggregated data, we could get data about each teacher, it would allow for greater insights and deeper analysis. If we were able to tie a teacher to a school, the analysis could potentially look at the outcomes for students in a school based on the teachers. Many states have policies against this analysis (judging teacher performance by student outcomes). For example,
    - It is difficult to track the extent to which trained teachers have moved to different schools or were unable to implement a CS course. The list of schools to which teachers are assigned is derived from either the application forms (most common) or through program records, which are updated by the project. Most projects confirm that the schools they provided to the EWG are schools in which the teacher is actively teaching CS. The EWG suspects, however, that a significant number of teachers may have moved schools, or are unable to implement a course despite best intentions (e.g., course may be cancelled to low enrollment, competing priorities at a school or personal problems), thus, the list likely has some schools in which a teacher is no longer teaching and/or include schools in which there is a teacher but the teacher is not implementing the course as planned.
    - It is unknown how many teachers participate in more than one PD opportunity. It is possible that a subset of teachers are prepared in both ECS and CSP through different PD providers.
- At present, there is no examination of or distinction between CS10K programs in terms of their quality. The current approach assumes all projects are equally efficacious for preparing teachers. While this is an important and ideally ultimate question for CS10K to examine, it also begs additional questions such as how one might define or measure effective training and CS instruction and the integrity of implementation. For example:
    - Evaluators have a difficult time collecting classroom implementation data and have long struggled to specify how professional learning changes classroom instruction. Currently, we ask how much professional learning material teachers use during classroom implementation. We acknowledge that this proxy for classroom implementation does not reveal day-to-day, program-related improvements to classroom instruction. These data do not account for what DeBarger et al. (2013) call productive adaptations (unintended improvements to instruction as a result of professional learning) nor program endurance across multiple years. These data also do not reveal program fidelity at two critical levels: structure (the framework for service delivery) and process (the ways in which the services are delivered) (See Mowbray et al. (2003) and Century & Casatta (2016).
- Reporting of race and ethnicity is problematic for interpretation. Our current approach in reporting race and ethnicity for teachers is challenging for some evaluators to provide and problematic for interpretation. Currently, the EWG asks about race and ethnicity as two distinct items as per the census structure and as recommended by the Office of Management and Budget. This, however, may force an individual that primarily identifies as Hispanic to also select a race category that may not be meaningful to them, and potentially overinflate the "White" response. These demographic questions also may differ from how evaluators have typically asked these questions. Student level data is obtained from administrative data;

thus, it is up to a school system to define the parameters. It also differs from how the NECS database stores race and ethnicity data.

State-level data is a feasible, sustainable option for tracking the impact of CS teacher PD

Using state-level data to help define broadening participation goals for computing education and for measuring progress allows states to understand the current landscape and track progress over time.

It is, however, difficult to prepare a request and to process data into meaningful knowledge for a variety of reasons including:

- Defining computer science and identifying courses
  - Difficulty defining Computer Science may be challenging for some states. When there are not clear standards associated with a CS curriculum, a specific definition is needed; one that is understood and shared by administrators, teachers, and other stakeholders
  - Identifying relevant course codes for inclusion in the data request and ensuring that these course codes accurately reflect what is happening in the classroom
  - Collecting historical data may be problematic due to changes in course names and numbers and may require additional course lookup.
- Difficulty engaging stakeholders to create a shared goal for broadening participation
- Limitations with the data systems themselves
  - Each state's system is unique and sometimes it requires accessing multiple systems to gather all relevant data
  - Courses may not be coded in a universal manner within a state
  - Courses that integrate CS rather than serve as a stand-alone CS course may or may not be identifiable
  - K-8 classes may be defined and tracked in a manner that precludes identifying CS systematically.
- Privacy concerns
  - May impede who can access the data and for what purposes
  - May have a threshold for the number of students at which they are able to provide information. This could potentially preclude accessing data from areas with small or emerging CS participation.
- Capacity challenges
  - Making the request requires a significant amount of time (see above challenges)
    - A formal request may not actually be the most effective means of requesting data. Political connections (that not all teams have) are important in the process.
  - Processing the data can be time consuming particularly if the data has errors, comes from multiple sources, or needs sophisticated analysis
  - Sharing data in a meaningful manner with specific stakeholder groups

Recommendations

Cross-project reporting is complicated but important. Establishing clear reporting requirements upon the issue of an RFP can minimize the confusion, burden, and resistance to participating in such an effort. Developing a clear set of variables that are common across projects can be generated from the program's theory of action which will subsequently allow for meaningful analysis of program impact.

Given the rapidly changing context of CS education, the EWG further recommends any program-wide data collection effort be supported by a contractor. There is an increasing investment on behalf of NSF (and others) in preparing teachers to offer CS courses. The EWG has defined a number of data elements that will continue to be collected from CS10K projects, but given the expansion of the CS PD landscape, it might be useful to consider a data collection system that goes beyond CS10K to include all projects tagged (or yet to be tagged) as CS For All (including CS10K, STEM+C and EHR). This system would ideally be flexible enough to allow project PIs or researchers to enter data (similar to the Math Science Partnership system) and also allow the project representatives (PI, manager, or evaluator for example) to enter data (similar to the system being used by EPSCoR). Though there is enormous goodwill within the evaluator community, the tasks of large scale data collection and compilation are burdensome and scaling up is likely to be very challenging. Requiring reporting through a dedicated system will streamline the process and improve the quality and quantity of data provided. Eventually this system may expand to include projects beyond CS10K (STEM+C and other EHR projects) and if robust enough, could include projects beyond NSF.

Lessons learned through CS10K cross-project data collection suggest that such an effort should:

- Identify the questions that are relevant to understanding the investment in a program
- Provide technical assistance and support to projects early in their implementation cycles
- Provide resources to work with district representatives to provide data about school and student participation

As CS10K project funding sunsets it would be worth understanding how the projects are sustaining their PD efforts (if at all). For projects that are continuing, how have they done so? For example, have they secured additional funding? Have districts or other organizations taken over the PD? If so, what have been the benefits and the pitfalls?

Appendix A: Memo: CS10K Close-Out Survey Feedback

Executive summary/abstract

In July 2018, a user-survey to capture feedback about the CS10K Common Data Collection effort was administered to all participants. The purpose was to address the following questions:

1. What was the burden (initial and ongoing) required to provide the data requested?
2. Which data elements were challenging to provide?
3. What EWG-provided resources did respondents use?
4. Were data elements integrated into respondents' evaluations or ancillary?
5. What recommendations, if any, did respondents have for NSF about collecting cross-project data?

Eleven participants responded, providing information about the process. Most participants found the burden to be reasonable and the data elements collected to be useful to their own evaluation efforts. Student data was more challenging to collect than teacher data due to IRB concerns and district policies.

Methods

In the summer of 2018 all PIs and evaluators that had participated in the CS10K Common Data Collection Effort were emailed a link to a "close-out" survey. The survey, designed to obtain information about the burden and value of the CS10K common data collection effort and tools can be found in Appendix E. Eleven PIs and/or evaluators responded to the survey; nine of these respondents completed the entire survey. Because all responses are anonymous, it is impossible to identify if the respondents represent unique projects or if multiple respondents provided information about a single project (for example, the evaluator and PI both responding to the survey). However, if only one respondent per project responded, the data represent feedback from approximately 25 percent of CS10K projects. Further, because responses were anonymous, we cannot determine to what extent the perspectives provided reflect those of PIs or evaluators. Nor do we know if the projects that elected to complete the survey are fully representative of the experiences on the other CS10K projects.

Findings

Burden

Respondents reported that participation in the CS10K common data collection effort required an average of approximately 11-12 hours and two staff persons. Respondents felt the time was worth the investment-- that the document providing suggested wording and the data shell were useful tools for evaluators. Comments included:

- *I got better in completing the shell because we knew to have data available. In general the summary data that the shell collects does not seem particularly sensitive to me and forces me to discard information on the individual level when I collapse counts over students.*

- *The process as a whole worked well. The data shell with the instructions was easy to use. The suggested question wordings were helpful and were used to make sure we were collecting the needed information*
- *The Excel spreadsheet shell was cumbersome to use, but that's really just an Excel issue. On the other hand, I do like the fact that it means I have a copy of all the data that I provided*.

Table 1 shows the extent to which respondents collected data as part of their own evaluation project versus solely for the EWG request. Written comments suggest that some respondents incorporated EWG wording or questions into their own evaluation and used this data independently while others were already collecting the data element as part of their independent evaluation.

Table 1. Extent to which respondents only collected data because of the EWG request and/or were able to use these items for the evaluation of their own project (N=11):

	Number of people who only collected item because of EWG request	Number of respondents who used item for evaluation of own project
Teacher gender	1	9
Teacher ethnicity	1	9
Teacher race	2	9
Teacher disability	7	4
Teacher years teaching K-12	1	6
Teacher years teaching Computer Science	1	9
Teaching HS CS courses: ECS, CSP, CS-A	5	9
Teaching course connected with PD	1	6
How are teachers teaching with PD	2	6
Endorsement, Certificate and/or Credentials	1	7
Primary discipline	0	7

**there was one response that either didn't answer at all or was a "no" for each category*

Data Elements

The CS10K data sheet asks for information about teachers' demographics, background, and implementation of PD; student demographics; and a list of schools with teachers who participated in CS10K-supported training.

Teachers

Table 2 presents each of the data elements for teacher information and the reported ease in providing this information.

Table 2. Respondents' assessment of the ease with which they were able to provide the requested data on teacher background, demographics, and implementation (N=11):

	Very difficult (1)	Some-what difficult (2)	Some-what easy (3)	Very easy (4)	Did not provide (n/a)	Don't know (n/a)	Aver-age
Teacher gender	0	0	4	7	0	0	3.6
Primary discipline (N=10)	0	0	3	6	1	0	3.7
Teacher ethnicity	0	0	5	6	0	0	3.5
Teacher race	0	0	5	6	0	0	3.5
Teacher years teaching K-12	0	1	3	7	0	0	3.5
Endorsement, Certificate and/or Credentials	0	0	5	4	1	1	3.4
Teacher disability	1	1	3	6	0	0	3.3
*Teacher years teaching Computer Science***	0	1	5	4	1	0	3.3
Teaching HS CS courses: ECS, CSP, CS-A	0	2	5	4	0	0	3.2
Teaching course connected with PD	0	3	3	3	1	1	3
How are teachers teaching with PD materials/approaches/and curriculum?	0	3	4	1	2	1	2.8

**did not include "did not provide" and "don't know" when calculating average **N=10 when calculating average*

Most of the information was easy to provide, on average. The highest average "ease" score was, somewhat surprisingly, the primary discipline in which a teacher is teaching (as this item was added in the 2015-16 data collection year). Less surprising were teacher demographics such as gender, race and ethnicity. The two most challenging items were if teachers offered the course connected to the PD and how they are using the materials from that PD. These two items required follow-up explanations. The participants explained:

- *[Project] already had ethnicity, race and disability so we had to ask teachers for information they'd already but the response options did not match the SageFox items so we asked for them for these data again.*
- *Teachers are reluctant to release disability data about themselves or their students, unless required to do so by their school/district. They most often default to "did not provide".*
- *We are able to collect the PD material usage, but it is difficult to determine the PD usage as our program allows for various PD materials and some teachers customized their own*

Teacher Disability Status

The EWG is particularly pleased to see the increased capacity of projects to provide disability data over time. Table 2 provides an overview of the percentage of projects that have been able to report on teacher demographics, and specifically on teacher disability status.

In 2015 we were unable to answer the question "How many teachers report having a disability as defined by the Americans with Disabilities Act" due to low reporting (it is unclear if projects asked the question, or if people declined to respond). In 2016 we were able to report on the 832 teachers for whom data was provided for this question. In 2017 there were data on 1,392 teachers. Now, in 2018, there are data on 2,063 teachers available (cumulative across 4 years of data collection). Of the teachers for whom data on disability status are provided, 92% reported no disability, 5% reported a disability and 3% preferred not to answer this question (See Table 3).

Table 3: Projects reporting disability status over time

	Total Projects Submitting Data	% Projects reporting teacher data	% Projects reporting teacher disability data
2014-2015	20	85%	45%
2015-2016	29	83%	59%
2016-2017	25	96%	72%
2017-2018	20	100%	70%

Student data

Gathering data about students was more difficult. The system requested demographic data for those students who (1) enrolled in a CS course; (2) passed a CS course; and/or (3) participated in the AP CSP exam (if relevant).

Similar to the teacher data, basic demographic items were the easiest to obtain (gender, race, and ethnicity), while information about disability status was more difficult to access. Further, respondents reported it was easier to gather information about students enrolled in a CS course than it was to compile data for students who passed a CS course or took the CSP AP exam (see Table 4)

Table 4 Please indicate the ease with which you were able to provide the requested data on student participation (N=11):

	Very difficult (1)	Somewhat difficult (2)	Somewhat easy (3)	Very easy (4)	Did not provide (n/a)	Don't know (n/a)	Aver-age
Gender of students enrolled in the course	0	5	0	4	2	0	2.9
Total number of students enrolled in the course	0	4	3	2	2	0	2.8
Ethnicity of students enrolled in the course	2	3	0	4	2	0	2.7
Race of students enrolled in the course	2	3	0	4	2	0	2.7
Total number of students who passed the course	1	4	1	1	3	1	2.3
Total number of students who took the AP exam	1	2	2	0	5	1	2.2
Gender of students who passed the course	1	3	0	1	5	1	2.2
Gender of students who took the AP exam	1	3	0	1	5	1	2.2
Ethnicity of students who passed the course	1	3	0	1	5	1	2.2
Ethnicity of students who took the AP exam	1	3	0	1	5	1	2.2
Race of students who passed the course (N=10)	1	3	0	1	5	1	2.2
Race of students who took the AP exam	1	3	0	1	5	1	2.2
Disability status of students enrolled in the course	4	2	0	2	3	0	2

Collection of student data is complicated by privacy concerns and district specifications for access to data. The EWG specifically requested all student data come from: a) direct student surveys for which a project had IRB approval or b) district records. For some projects, the cost of gathering this data was prohibitive given their evaluation budgets. Comments concerning these themes include:

- *School districts refuse to provide information on disability because of privacy concerns. School districts will often not provide data on gender and race.*
- *We had no follow-up ability to track which students passed the class, took the AP exam, etc.*
- *We did not have IRB approval to gather student data. We were reliant upon teachers to provide answers to the question is about students that we asked.*
- *These data were not systematically tracked by the project, though they should have been!*

- *We have worked with large numbers of students. These demographic items were asked students on a pre-survey, but not all students got the pre-survey and some who did take the pre-survey did not stay in the class. Sorting data out from the student pre-survey was very difficult. For the last several years we also asked for student demographic information from teachers at the end of the course. These data were easier to deal with, but we did not always get teachers completing the survey on which this information was collected. Data on whether students took the AP exam are always difficult, it large part because we do not get individualized student results on the AP.*
- *It took several layers of communication and wait time to obtain district approval to release requested student data*

Those who collected the student data elements did so as part of their evaluation with the exception of disability status.

Discussion

As reported, survey respondents reported that the CS10K Common Data Collection Effort was worth the burden, which was not onerous. According to one respondent:

- *The only real hindrance to data collection was the project's limited evaluation budget. We were unable to use multiple methods for gathering data from teachers many of whom we had lost contact with in the third year: We used only online surveys and announcements on the forum, rather drawing teachers in that we could not reach through email with snail mail or other methods.*

Respondents to the survey noted that the support offered by the EWG was an important part of the success of this effort both for facilitating consistent reporting but also for providing data elements that were useful to the individual project evaluations.

Moving forward, survey respondents report that the CS10K Common Data Collection Effort is "a good example of how to collect cross-program data" but that such an effort needs to start earlier in the process. One respondent commented that:

- *I think an organization like SageFox should get a contract before the projects are funded so that evaluators aren't learning about the data collection requirements after they are funded. In this case it was painless but evaluators need to know what they'll be expected to do before they propose (and budget) an evaluation plan.*

Finally, some participants had hoped such a cross-project data collection could do more to measure longitudinal outcomes and impacts.

Appendix B

Definition of Teacher Participants

For data collection purposes, projects were asked to define teacher participants as those who are active on an annual basis, starting in June (roughly) with summer professional learning and teaching through the following academic year. Participants recruited during the spring are, therefore, not included in any common data collection until they have (typically) begun intensive PD during the summer months preceding the academic year in which they are to actually implement the CS10K training. This is especially relevant to projects funded in 2016, as most were structured to begin offering PD to participants starting summer of 2017, and into the 2017 - 2018 school year. We present our results in 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 2013-14 refers to data collected from projects about teachers who began their PD during summer 2013 and may have taught CS through the 2013-14 school year.

The Common Measures and Reporting Tool

The 2016-17 data shell is largely consistent with the previous year's version (SageFox 2016) with the exception of participant tracking for the AP CSP exam. Specific tabs of the data collection spreadsheet include:

1. Teacher descriptives/demographics
    a. Demographics: gender, race, ethnicity, disability status
    b. Experience teaching both in a K-12 school and CS
2. Teacher implementation
    a. Use of PD
    b. Credentials held and primary teaching field
3. Student-level data
    a. Student enrollment (enrollment and pass rates)
    b. Student demographics: gender, race, ethnicity, disability status
    c. Participation in the AP CSP exam
4. School data
    a. List of schools in which CS10K teachers are teaching

The EWG created a technical document to accompany the common elements data shell. In an effort to facilitate aggregation across projects (as well as within projects over time), the document it contains survey items that the EWG encouraged CS10K evaluators use in their teacher surveys. This was especially critical for demographic items designed to obtain race/ethnicity and disability status data about teachers, which can be particularly sensitive, and thus challenging to collect. In our recommended survey items that ask about race/ethnicity, 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 (for example, if an individual selected both Asian and Black, they would be counted in each category). This approach might also overinflate the number of teachers identified as "White" as a person who considers him/herself Hispanic, may have felt forced to select White as a race, when race may not be a defining characteristic for that individual. Collecting student-level data of this nature requires extensive work with district IRB/RRB offices and

thus projects were instructed only to provide this information if they had such approval, likely leading to a significant undercounting of students in this report.

Data Collection Process

EWG team provided projects with a blank data shell and "technical documents" with recommended survey items as a companion to the data shell. The team also assigned a EWG liaisons to a small number of evaluators of CS10K projects and address any issues they might encounter. In mid-spring 2018, we sent evaluators the EWG Data Shell and the technical document in preparation for their 2018 end-of-school-year survey administration

The EWG continued a liaison approach to data collection. The 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. The liaisons sent each project their data shell template, technical support document, and the year 2 report. EWG liaisons followed-up with project evaluators throughout the spring 2018 to prepare them for accurate and timely data submissions.

Caveats

The common data elements shell prompted evaluators for any data caveats so we could better understand the collection challenges. Using this information and the findings from the data validation interviews (conducted in spring 2017) we discovered the following common data collection challenges:

- **Some projects don't fit neatly into the data shell structure** due to the way they implement their professional development (e.g. CS CAVE, IDOcode) which may have a different cycle of implementation than the academic year and may not have the same relationship with teachers who are trained.
- **Projects used different methods to obtain information about teacher demographics**. Although evaluators have been able to consistently provide data about teachers' gender, race and ethnicity over the three data collection cycles, they used different methods to obtain this information. Typically, projects used teacher surveys to capture this information, but occasionally used application or program records. This may affect interpretation of the data reported.
- **Projects used different timeframes to report teacher experience.** Most projects asked for the participants' teaching experience for both K-12 and CS courses. The data provided to the EWG seem to be evenly split between projects that collect these data on applications or at the start of PD projects and projects that collect these data at the end of the implementation year. This means that the data for some projects may be "off" by a year (for example, a teacher may have 10 years of K-12 experience at the point of entry during Spring 2015 but the data collected by the EWG would ask for their experience *including* the implementation year, which would be 11 years by Spring of 2016).
- **Projects used different methods to track teacher implementation.** In addition, evaluators were asked if they collected implementation data beyond the first training year. About half of the projects interviewed report that they do track teacher implementation for the duration of the grant. Several reported that it is challenging to do so when the teachers have no incentive to participate as their formal obligation to the program is over (typically after one year). The data collection spreadsheet currently asks evaluators to report on previously trained teachers, but for some projects this is not possible or there are large amounts of missing data.
- **Projects use different methods to collect information about teacher implementation.** During data validation interviews in Year 2, the EWG learned that the two most common

methods were teacher surveys administered at the end of the year or at the end of each unit. In the former, the question used was typically that which was recommended by the EWG. In the latter, the evaluator may have retrofitted data from observation, teacher report (e.g interviews or surveys not using the EWG recommended questions) or other records to fit the EWG common data spreadsheet. This leaves room for evaluator interpretation that may be inconsistent across projects. It is unknown to what extent this remains true for Years 3 and 4.

- **Projects made different interpretations of the requested respondent pool for teacher implementation.** During data validation interviews in Year 2, the EWG asked projects regarding whom they administered the teacher implementation questions. The EWG intended data only to be provided for teachers implementing. In most cases the projects that were contacted reported an accurate interpretation, however in at least one case the project had included teachers who are not implementing and reported they were using "none" of the material. We also learned that some evaluators used the data they collected independently through interviews and/or unit surveys to retrofit the data to the EWG question format. Given the challenges with interpretation and the low response rate, the response to this question may not be truly representative. It is unknown to what extent this remains true for Years 3 and 4.
- **Projects struggle to collect student participation data. S**tudent data are perhaps the trickiest data to collect given IRB/RRB requirements. For the projects that have successfully collected data, it is typically done through district records and/or direct student surveys. Data provided to EWG via direct student surveys are limited to those for which there is parental consent/student assent; therefore, the data provided is often for a subset of the students reached. There are several projects (exact number unknown) that collect the student data through teacher surveys (although teachers are asked not to "guess", but to draw student demographics from administrative records).
- **There are more CSP projects and teachers in the overall data set, but the students are equally represented between CSP and ECS** with about 13,000 students participating in each course**.** It is important to note that there are significantly more CSP projects and teachers in the overall dataset, and yet there are more students in the ECS projects that have been accounted for. This may be due to better and more systematic data collection than the CSP projects have the capacity for, smaller classes, or the "newness" of the course. Regardless, the diversity of students should be interpreted with extreme caution.
- **ECS is typically concentrated in major urban areas, unlike CSP which is a national offering.** ECS in only 11 states and largely concentrated in Urban districts. Thus, the available student data may be overly representing specific regions and not representative of all students reached through CS10K.
- **It is difficult to track the extent to which trained teachers have moved to different schools or were unable to implement a CS course.** The list of schools to which teachers are assigned is derived from either the application forms (most common) or through program records, which are updated by the project. Most projects confirm that the schools they provided to the EWG are schools in which the teacher is actively teaching CS. The EWG suspects, however, that a significant number of teachers may have moved schools, or are unable to implement a course despite best intentions (e.g., course may be cancelled to low enrollment, competing priorities at a school or personal problems), thus, the list likely has some schools in which a teacher is no longer teaching and/or include schools in which there is a teacher but the teacher is not implementing the course as planned.
- **It is unknown how many teachers participate in more than one PD opportunity.** It is possible that a subset of teachers are prepared in both ECS and CSP through different PD providers.

Appendix C: Student Demographics

		2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
	Total number of course sections for which student data are reported			25	98	306	512	551
	Total number of course sections taught by project teachers	67	57	105	146	268	444	483
	Total number of teachers for whom student data is provided	55	68	170	242	380	232	459
	Total students enrolled in all sections of course instructor taught related to training (ECS or CSP)	3099	3858	7019	11013	19140	15525	23,708
Student Gender	80. Student Gender - Female	1146	1505	2206	2873	4431	3877	8780
	81. Student Gender - Male	1626	2176	2803	5167	9080	8395	13163
	82. Prefer Not to Answer	0	5	4	7	22	42	10
	83. Data Not Available	0	0	146	566	215	710	453
Student Ethnicity	84. Student Ethnicity (Hispanic) - Yes	1916	2244	2577	2970	3591	2594	6488
	85. Student Ethnicity (Hispanic) - No	856	1440	2435	5032	8561	5179	7943
	86. Prefer Not to Answer	0	0	0	9	22	696	54
	87. Data Not Available	0	2	147	602	1519	1495	1401
Student Race	88. American Indian or Alaska Native	1	43	115	123	159	93	132
	89. Asian	182	397	410	563	1016	1058	1227
	90. Black or African American	251	669	478	1408	3043	1063	3870
	91. Native Hawaiian or Other Pacific Islander	5	6	52	43	210	42	38
	92. White	97	350	1059	2870	3829	6768	6305
	93. More than one race selected	7	17	71	267	181	645	390
	94. Prefer Not to Answer	0	0	61	53	136	810	87
	95. Data Not Available	72	115	589	753	1626	1032	7502
Student Disability Status	96. Yes	0	0	0	119	281	379	1683
	97. No	0	0	0	260	2960	5990	11770
	98. Prefer Not to Answer	0	0	0	0	36	844	15
	99. Data Not Available	2772	3686	4615	5289	7407	1444	2634

Appendix D

Full Spreadsheet found here:
https://docs.google.com/spreadsheets/d/1YZeH5NusEG12kNcDTwAcjPO0I4DMtSHKugg9lAHooNk/edit#gid=1060482857

Contact Information

Evaluator Name	Evaluator Email Address	EWG Project Liaison Name	EWG Project Liaison E-mail	Evaluator Phone	Project Name	NSF Award #	CS10K project focus (CSP, ECS, other)	Time required to complete this submission (in hours)	Caveats / additional information unique to this submission
John Q. Beancounter	Example_Eval@email.com	Jeff Xavier	Example_liaison@email.com	555-3924	BJC4NYC fake				

Evaluator Name	Evaluator Email Address	EWG Project Liaison Name	EWG Project Liaison E-mail	Evaluator Phone	Project Name	NSF Award #	CS10K project focus (CSP, ECS, other)	Time required to complete this submission (in hours)	Caveats / additional information unique to this submission
John Q. Beancounter	Example_Eval@email.com	Jeff Xavier	Example_liaison@email.com	555-3924	BJC4NYC fake				

Instructions:
- Please complete one Excel workbook per CS10k project.
--Please provide all data available for each sheet. If there is missing data (or data you are unable to share due to IRB concerns) please provide a note in the 'caveats/notes' section on each page

Submission of data:
- A final draft of the data should be submitted on [Date].
- Please submit your work to cs10k@sagefoxgroup.com and CC your liaison
- The technical document containing additional information about the survey questions upon which this sheet was created can be accessed by clicking anywhere in this text box.

Thank you very much for your help!

BJC4NYC fake		Basic Demographics																			
	Number of teachers successfully completed training	Gender (TD1)				Ethnicity (Hispanic/Latinx) (TD2)				Race (TD3)								Disability (Y/N) (TD4)			
	Options: number or x (meaning NA)	Female	Male	Prefer Not to Answer	Data Not Available	Yes	No	Prefer Not to Answer	Data Not Available	American Indian or Alaska Native	Asian	Black or African American	Native Hawaiian or Other Pacific Islander	White	Other	Prefer Not to Answer	Data Not Available	Yes	No	Prefer Not to Answer	Data Not Available
Newly trained (2016-2017)	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#

Teacing History																					
Number of years taught K-12, including the 2016-2017 academic year (TD5)								Number of years taught CS at the high school level, including the 2016-2017 academic year (TD6).							Prior experience teaching CS courses (TD7)						
Not yet teaching in K-12	1 Year	2-3 Years	4-5 Years	6-10 Years	11-15 Years	16+ Years	Data Not Available	Not yet teaching CS	1 Year	2-3 Years	4-5 Years	6-10 years	11-15 years	16+ Years	Data Not Available	# with ECS experience	# with CSP experience	# with AP CS A experience	No prior experience teaching these courses	Data not available	Comments, caveats and additional information to help us understand if there were special circumstances or problems with certain data that we should understand
#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	

Teacher Background & Demographics

Instructions:

1. In row 4, indicate the number of teachers who participated in your CS10K project for the first time during the 2016-17 academic year.
2. Column B should contain either a number or "X" (X indicates that this row is not applicable for your program). If you enter a number, all cells in that row will turn red until they are populated. If you enter an "X" all fields will turn gray to indicate that they are not applicable.
3. Fields with "#" should contain only numbers.
4. Each set of columns highlighted in an individual color represents a single "variable" (e.g. gender is highlighted in light blue). The numbers of teachers provided across the different response options for each individual variable should sum to the same number as that in column B.
5. Please use teacher-provided data to complete gender, ethnicity, race, and disability status. Please do not complete this based on teacher names or your perception of these demographic items. If you do not have this information for a teacher, count them in the "Data Not Available" cell.
6. If you used different categories than those provided for race/ethnicity, please provide your data as best fits. Please make use of the 'other' option as appropriate.
7. If you used different categories for years teaching please provide your data as best fits. Please use the 'caveats' column for any explanations.
8. Successful completion of training is up to the project to define. In general, it would mean completing the full offering of professional development but may or may not be teaching in the classroom yet
9. PLEASE NOTE: If data for any field or set of fields were not available due to issues with getting IRB approval to collect the data, please discuss this in the comments cell (column AR).

Instructions:

1. In row 4, indicate the number of teachers who participated in your CS10K project for the first time during the 2016-17 academic year.
2. Column B should contain either a number or "X" (X indicates that this row is not applicable for your program). If you enter a number, all cells in that row will turn red until they are populated. If you enter an "X" all fields will turn gray to indicate that they are not applicable.
3. Fields with "#" should contain only numbers.
4. Each set of columns highlighted in an individual color represents a single "variable" (e.g. gender is highlighted in light blue). The numbers of teachers provided across the different response options for each individual variable should sum to the same number as that in column B.
5. Please use teacher-provided data to complete gender, ethnicity, race, and disability status. Please do not complete this based on teacher names or your perception of these demographic items. If you do not have this information for a teacher, count them in the "Data Not Available" cell.
6. If you used different categories than those provided for race/ethnicity, please provide your data as best fits. Please make use of the 'other' option as appropriate.
7. If you used different categories for years teaching please provide your data as best fits. Please use the 'caveats' column for any explanations.
8. Successful completion of training is up to the project to define. In general, it would mean completing the full offering of professional development but may or may not be teaching in the classroom yet
9. PLEASE NOTE: If data for any field or set of fields were not available due to issues with getting IRB approval to collect the data, please discuss this in the comments cell (column AR).

Implementation

BJC4NYC fake																						
New vs. Returning	Number of teachers who successfully completed training	Teaching of course connected with PD (T11)			Use of PD materials in course for which teacher received training (T12)							For which subjects does teacher have an endorsement or credential? (Choose all that apply)										
	Options: number or x (meaning NA)	Yes (teaching course this year)	No (not teaching course this year)	Data Not Available	Teach using only materials from training	Teach primarily with materials from training	Teach about half with training materials and half with other	Teach primarily with other materials, use training materials to supplement	Do not use any training materials in teaching course	Did not teach course associated with training	Data Not Available	Comp. Science	Math	Science	Business	Career or tech. ed.	English / Social Stud. / Human. / Other	Other credential	No credential / not currently teaching	Data not available
Newly trained (2016-2017)	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#
Previously trained (implementing in 2016-2017)	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#

Instructions:

1. On this sheet teachers are divided into those who are newly trained - i.e. having participated in the project for the first time in the 2016-2017 academic year - versus those who first participated in the project in prior academic years but are still being tracked by the evaluation.
2. Fields with a "#" should contain only numbers.
3. PLEASE NOTE: If data for any field or set of fields were not available due to issues with getting IRB approval to collect the data, please discuss this in the caveats section (column AE)

For which subjects does teacher teach more than 50% of their courses? (Primary Discipline)										
Comp. Science	Math	Science	Business	Career or tech. ed.	English / Social Stud. / Human. / Other	Other discipline	Teach two or more disciplines equally	Not currently teaching	Data not available	Comments, caveats and additional information to help us understand if there were special circumstances or problems with certain data that we should understand
#	#	#	#	#	#	#	#	#	#	
#	#	#	#	#	#	#	#	#	#	

Student Participation

Instructions:

1. This sheet asks for demographic data on the students taught by all project teachers currently implementing training material in a relevant course (including those who were trained this project year as well as those trained in prior years). This includes students in 3 categories:
   - All students enrolled in course
   - All students who passed course
   - All students who took AP CSP exam
2. In order to have a better understanding of missing data, we ask that you report three items: 1) The total number of teachers for whom student data was obtained 2) The total number of course sections taught by ALL teachers involved in the project in a given academic year and 3) The total number of course sections for which student data was obtained
3. Fields with a "#" should contain only numbers, or "x" or "X" to indicate data not available.
4. If you used different categories than those provided for race/ethnicity, please provide your data as best fits. Please make use of the 'other' option as appropriate.
5. PLEASE NOTE: If data for any field or set of fields were not available due to issues with getting IRB approval to collect the data, please discuss this in the caveats section (column Z)

BJC4NYC fake						Student Gender (S2)				Student Ethnicity (Hispanic) (S3)			
	Total number of course sections taught by project teachers	Total number of teachers for whom student data is provided	Total number of course sections for which student data are reported	Student Category	Total Number of Students	Female	Male	Prefer Not to Answer	Data Not Available	Yes	No	Prefer Not to Answer	Data Not Available
Students reached in 2016-2017 academic year	#	#	#	Enrolled in course	#	#	#	#	#	#	#	#	#
				Passed Course	#	#	#	#	#	#	#	#	#
				Took AP CSP exam	#	#	#	#	#	#	#	#	#

Student Race (S4)								Student Disability Status (S5)				
American Indian or Alaska Native	Asian	Black or African American	Native Hawaiian or Other Pacific Islander	White	More than one race selected	Prefer Not to Answer	Data Not Available	Yes	No	Prefer Not to Answer	Data Not Available	Comments, caveats and additional information to help us understand if there were special circumstances or problems with certain data that we should understand
#	#	#	#	#	#	#	#	#	#	#	#	
#	#	#	#	#	#	#	#	#	#	#	#	
#	#	#	#	#	#	#	#	#	#	#	#	

School Information

Instructions:
[This text box may be deleted]

1. This sheet asks for key information about each of the schools that have been reached through the CS10K project. Any schools that have been previously submitted have been pre-populated on your spreadsheet. You may add as many additional schools as you like.
2. For each known school, supply the school's name, zip code, district name and whether it is public or private.
3. For each school, provide in column E the number of unique teachers who teach at the school, but are not currently (academic year 2016-17) teaching the CS course for which they were prepared
4. For each school, provide in column F, the number of unique teachers who teach at the school and are currently (academic year 2016-17) teaching the CS course for which they were prepared.
5. For each school, provide in column G, the number of unique teachers who teach at the school but you do not know if they are actively (academic year 2016-17) teaching the CS course for which they were prepared.
4. Any difficulties or anomalies related to these data can be reported in the "caveats" column (column I)

BJC4NYC fake								
School Name	District Name	Zip Code	Public/Private	Number of teachers trained in 2016-2017 academic year who used CS10k training materials in their teaching	Number of teachers trained in 2016-2017 academic year who teach at this school but are not using CS10k training materials in their teaching	Number of teachers trained in 2016-2017 academic year for whom use of CS10k training materials in teaching is unknown	NCES Code	Comments, caveats and additional information to help us understand if there were special circumstances or problems with certain data that we should understand
Example school	District 9	12345	Public	#	#	#		

Reporting and Dissemination

The following procedures have been approved for the CS10k Evaluator Working Group by the UMass Amherst Institutional Review Board**

As part of the ECEP project a supplement to the ECEP grant has been funded by the National Science Foundation (NSF). The work sponsored under this portion of the grant focuses on collecting information about teachers and students sponsored under the NSF CS10k initiative, which is aimed at increasing the number of computer science teachers nationwide. Work being done under this phase of the project (henceforth referred to as the CS10k supplement) mostly involves collecting aggregated data from evaluators of CS10k grants about their participating schools, teachers and students. The collection and dissemination of data collected under this portion of the study will have been approved by the IRBs under which each of the CS10k evaluators operates, and therefore will already be in compliance with human subjects requirements.

The activities specific to the CS10k supplement are intended to obtain a high level overview of school, teacher and student participation in the 30+ NSF-sponsored CS10k professional development projects currently funded across the country. The study aims to understand how many teachers are participating (including some basic background and demographic information about them), the extent to which they are making use of the material from their professional development, the schools at which they are teaching and the aggregate number of students (including some basic demographic information) they are teaching computer science to.

The CS10k Data Collection Shell spreadsheets will be identifiable by project and evaluator, but will remain confidential and will not be shared beyond the immediate members of the Evaluator Working Group. All responses reported back to the broader CS10k community will be aggregated so that it will not be possible to directly trace them back to individual evaluators. Data reported to the National Science Foundation may be broken down by project, such that it will be possible to determine what information was reported by an individual project. The high-level nature of the data being collected through this project will ensure that no data collected through this effort can be traced back to any individual school, teacher or student.

Summary results (including results from each CS10K project evaluator's data sheet submission) will be submitted to the National Science Foundation (NSF).

Articles may be submitted to conferences and/or peer-reviewed journals. Data will be presented in the same way that it is presented to NSF. NSF may choose to disseminate the CS10K project findings we compile through sharing reports or other dissemination avenues, which may be online or in print, with the goal of providing information about the total number of and background/demographics of schools/teachers/students impacted by the NSF CS10K initiative.

Data collected for the ECEP project, including the CS10K Supplement will follow the following procedures where applicable: All original versions of the paper-based surveys will be kept in a cabinet kept inside a locked room either at the ECEP program office at UMass Amherst or at SageFox Consulting Group and will be destroyed after the end of the project. The responses will be compiled into a database that is kept on a password protected server owned by SageFox Consulting Group. The online surveys and collected course data will be transferred directly to data files kept on the same server. Only evaluators at SageFox Consulting Group will have direct access to these data. All data downloaded from the online server onto an outside computer will either be password protected or re-uploaded after use and deleted from the computer.

All data collected through the CS10k effort will be electronic. These data will be stored on a password protected server owned by SageFox Consulting Group.

**Note: The above text is extracted from the larger ECEP IRB to enhance brevity and readability. To obtain the full text of the IRB submission, please contact your liaison.

(2017-2018 Academic year only)
Each section of the data collection spreadsheet includes instructions. This document is designed to provide specific guidance for data collection through surveys or other records. Each section references one of the tabs included in the data collection spreadsheet. These questions are designed to be administered at the END of the 2017-2018 academic year. Each tab has a "comments, caveats and additional information" column in which you can explain any nuances in your data or special consideration that should be used when interpreting it.

Contact Information Tab

This tab is the cover sheet to your data submission form. In this tab we ask for your name and contact information.

I. Teacher Background and Demographics Tab

This section requests data that may be gathered through survey data. All data should be collected from the participating teachers and not based on project personnel's "guess." The following are a list of survey questions to gather this data from participants. These questions may be/have been gathered as part of the program application and/or through evaluation surveys.

Item TD1: Gender

Item Stem: What is your gender identity?
Response Options:

- Female
- Male
- Prefer not to answer

Item TD2: Ethnicity

Item Stem: Do you identify as Hispanic or Latinx?
Response Options:

- Yes
- No
- Prefer not to answer

Item TD3: Race

Item Stem: Which of the following categories describes your race? [Select all that apply.
Response Options:

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White
- Other
- Prefer not to answer

Item TD4: Disability (Y/N)

Item Stem: Do you have a disability as defined by the Americans with Disabilities Act? (Under the Americans with Disabilities Act, an individual is considered to have a disability if he or she has a physical or mental impairment that substantially limits one or more of his or her major life activities, has a record of such impairment, or is regarded as having such an impairment.)
Response Options:

- Yes
- No
- Prefer not to answer

Item TD5: Number who replied that they had taught K-12 for the given number of years

Note: The response options below are how the EWG will ask for this information. For purposes of project evaluation, you may elect to ask for a specific number of years for which you would then recode into the options below when aggregating data for the common data submission.

Item Stem: Including this current school year, for how many years have you taught at the K-12 level?

Response Options:

- I have not yet taught in a K-12 school.
- 1 year (this is my first year)
- 2-3 years
- 4-5 years
- 6-10 years
- 11-15 years
- 16+ years

Item TD6: Number who replied they had taught *Computer Science* for the given number of years, including the current year

Note: The response options below are how the EWG will ask for this information. For purposes of project evaluation, you may elect to ask for a specific number of years for which you would then recode into the options below when aggregating data for the common data submission.

Item Stem: Including this year, for how many years have you taught a core Computer Science course, (such as Exploring Computer Science; Computer Science Principles or AP Computer Science A) at the high school level?

Response Options:

- I have not yet taught a Computer Science Course at the high school level
- 1 year (this is my first year)
- 2-3 years
- 4-5 years
- 6-10 years
- 11-15 years
- 16+ years

Item TD7: Experience with teaching CS

Item Stem: *Prior to the 2016-2017 academic year, had* you taught any of the following courses? *(Select all that apply)*

Response options:

- Exploring Computer Science (ECS)
- Computer Science Principles (CSP)
- Advanced Placement Computer Science - A (AP CS-A)
- No prior experience teaching these courses

II. Teacher Implementation Tab

This section requests data that may be gathered through survey data unless otherwise noted. All data should be collected from the participating teachers and not based on project personnel's "guess." The following are a list of survey questions to gather this data from participants. These questions may be/have been gathered as part of the program records and/or through evaluation surveys.

Item TI1: Teaching of course connected with PD

This item may be collected through administrative records

Item Stem: This current school year, are you teaching the computer science course for which you received professional development (e.g., Exploring Computer Science or Computer Science Principles)?

Response Options:

- o Yes, I am teaching the computer science course for which I received professional development
- o No, I am not teaching the computer science course for which I received professional development

Item TI2: How are teachers teaching with PD materials/approaches/and curriculum? (Question should only be asked if answer to TI1 was "yes")

Item Stem: For the computer science course for which you received professional development, to what extent are you teaching with instructional materials and /or approaches provided in the [ECS/CSP] professional development sessions?

Response Options:

- o I teach using only [ECS/CSP] materials
- o I teach primarily with [ECS/CSP] materials along with a few other supplementary materials
- o I teach with about half (50%) [ECS/CSP] materials and half (50%) other materials
- o I teach primarily with other materials and only use [ECS/CSP] to supplement my other, primary materials
- o I don't use [ECS/CSP] materials at all in my teaching.

Item TI3: Endorsement, Certificate and/or Credentials

Item Stem: In which areas do you currently hold an endorsement, certificate or other credential to teach? [check all that apply]

Response Options:

- o Computer Science
- o Math
- o Science
- o Business
- o Career or Technical Education courses
- o English/Social Studies/Humanities
- o Other
- o I do not have a certificate or credential to teach in a specific discipline

Item TI4: Primary Discipline

Item Stem: What is your primary teaching discipline? (the subject area in which you currently teach 50% or more of your courses?)

Response Options:

- o Computer Science
- o Math
- o Science
- o Business
- o Career or Technical Education courses
- o English/Social Studies/Humanities
- o I currently teach in two or more disciplines for an equal amount of time
- o I'm not currently teaching
- o Other

III. Student Participation Tab

Instructions**:** This section should be answered in compliance with the IRB/RRB for your project. Student may be collected through teacher surveys if in compliance with an IRB. All other questions should be collected through student surveys or school records. Teachers may not "guess" student demographic data.

To help contextualize the data, we are asking for:

- The total number of course sections taught by all teachers in the project
- The total number of teachers for whom student data is provided
- The total number of course sections for which student data are being reported

The data on this tab should be parsed in the following ways:

1. The total number of students enrolled in the courses for which the teacher was prepared to teach
2. The total number of students who passed the courses, as defined by the school or district.
3. For CSP courses, the total number of students who have taken the AP CSP exam.

Item S1a: Total number of students enrolled (For inclusion on teacher surveys)
Item Stem: How many students completed or are expected to complete the [ECS/CSP] courses you taught this school year?
Response Options: [Open numerical text]

Item S1b: Total number of students passed (For inclusion on teacher surveys)
Item Stem: How many students passed, as defined by your school, the [ECS/CSP] courses you taught this school year?
Response Options: [Open numerical text]

Item S1c: Total number of students participating in the AP CSP exam (For inclusion on teacher surveys)
Item Stem: How many students took the AP CSP exam this year?
Response Options: [Open numerical text]

Item S2: Gender (For inclusion on student surveys)
Item Stem: What is your gender identity?
Response Options:

- Female
- Male
- Prefer not to answer

Item S3: Ethnicity (For inclusion on student surveys)
Item Stem: Do you identify as Hispanic or Latinx?
Response Options:

- Yes
- No
- Prefer not to answer

Item S4: Race (For inclusion on student surveys)
Item Stem: Which of the following categories describes your race? [select all that apply]
Response Options:

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White
- Prefer not to answer

Item S5: Disability Status (For inclusion on student surveys)

Item Stem: Are you identified as a student with a disability by having either an Individualized Education Program (IEP) or a Section 504 Plan?
Response Options:

- Yes
- No

o Prefer not to answer

IV.Schools Tab

Item Sc1: Schools in 2015-2016 in which teachers are teaching CS

instructions: To collect the number and location of each school benefiting from CS10K-trained teachers, we need to know the school name, district name, zip code and if it is public private. This information may be collected in a variety of ways (for example survey of teachers or project administrative records). In this tab you will find a list of the schools previously submitted by your project.

For each school, please indicate the presence of teachers in each school for the 2016-2017 academic year by providing:

- Total number of teachers who are teaching who are using the CS10K training materials in their teaching
- Total number of teachers at each school that are NOT using the CS10K training materials
- Total number of teachers for whom use of CS10K training materials is unknown

If possible, it would be best to examine teachers' school data in the context of their response to questions about implementation of training materials (e.g. items TI1 and/or TI2) in order to place individuals in the proper one of these three categories.
Below is an example of a survey question you may use.

Item Stem: For each school in which you teach, please provide the following items:

Response options:

School 1: School Name
School 1: District name
School 1: Zip Code
School 1: Public/Private [select one]

School 2: School Name
School 2: District name
School 2: Zip code
School 2: Public/Private [select one]

School 3: School name
School 3 District name
School 3: Zip Code
School 3: Public/Private [select one]

Appendix F: Methods for the assessment of state data capabilities

In 2017 the EWG partnered with ECEP more deeply to better understand the opportunities and challenges with collecting statewide data. The EWG undertook three tasks:

1. Interviewing a subset of individuals in ECEP states to understand the procedures associated with making statewide data requests
2. Developing a survey to measure the *ease* and the *value* of defining measurable goals for broadening participation and to measure the *ease* and *value* for the collection of state-level data about courses, teachers and students.

3. Providing structured exercises for ECEP state teams to discuss the measurement of broadening participation in computing in their states.

State data interviews

Six interviews with state data representatives were conducted and analyzed. The states were Connecticut, Indiana, Massachusetts, Maryland, South Carolina, Texas, and Utah. The EWG also conducted informal interviews with several others in 2017 to understand the feasibility of preparing a state-wide data request. The ultimate objective was to enhance our capacity to examine and document the influence of PD through funding from NSF CS10K on teachers and their students. Interviews lasted approximately 45 minutes and looked to understand how researchers might:

1. Access state-held data including the process for making a request, the type of information available and for what time periods and the best time of year in which to make a request.
2. Analyze data, including ability to identify CS courses, grade spans, dual-enrollment courses, the level at which data is offered (individual, school, district or state), and the format in which data would be received.
3. Adhere to conditions for reporting or publishing any findings.

Ease x Value Surveys

The ECEP and EWG teams developed two surveys in advance of the January 2018 ECEP Annual Summit,.one about broadening participation (the BPC Survey) and one about state data collection (the State Data Collection Template). Both surveys were designed to capture the ease and value of data collection. In advance of the meeting, states participated on a November 16th, 2017 webinar with Daryl Chubin and Rebecca Zarch to discuss the need for rigorous data systems that allow for tracking broadening participation. Chubin presented the results of an NSF-funded workshop titled: ***Better STEM Outcomes: Developing an Accountability System for Broadening Participation*** and Zarch provided an overview of the Ease/Value templates and how they tied into the work of Chubin and his colleagues. Participants were tasked with completing the templates within 6 weeks so they could be used to inform the annual meeting. They were NOT expected to make a data request or have data in hand for the meeting. The two instruments are described below.

The BPC survey was developed largely by the ECEP leads and served as a precursor to the State Data Collection survey. Relevant fields from the BPC survey include:

- Defining the BPC work including definitions of
  - High quality CS
  - What is meant by CS for All
- Goal setting
  - Defining success in BPC
  - Communicating effectively with data
- Measurement
  - Identifying baseline data
  - How they will measure BPC
  - Tracking inclusion, retention and diversity
- CS Pathways
  - How to define pathways
  - How to measure achievement and BPC in pathways

States were asked to provide an indication using a 4-point likert scale of how *easy* and how *valuable* it would be to have these items defined:

Ease	Value
3 = Defining this information is easy for our state team	3 = This information is highly valuable to our state team

2 = Defining this information is somewhat difficult/complex for our state team 1 = Defining this information is very difficult/complex for our state team 0 = I don't know	2 = This information is somewhat valuable to our state team 1 = This information is not valuable to our state team 0 = I don't know

The State Data Collection Template was developed by the EWG with the intention that each state team indicate how *easy* and how *valuable* it would be to collect different types of institutional data about the individual, the school the district and the state. The template asked for the teams to identify how easy it would be to gather this data for all CS as defined by the state, but also for CSP and ECS courses specifically. The core of the template included four questions:

1. What is being taught?
2. Who is teaching CS?
3. Who are (and are not) the students taking and completing a CS course?
4. How well do students perform in CS courses?

The full set of survey items is below:

Core questions	Sub questions	Level			
		Ind.	School	Dist.	State
What is being taught (All CS; breakout by CSP, ECS)?	How many sections of each course?	E/V	E/V	E/V	E/V
	What are the characteristics of the course? 1. Grade bands 2. Are any of these blended courses? 3. Is the course part of a pathway? 4. Does the course count as dual enrollment?	E/V	E/V	E/V	E/V
Who is teaching CS (All CS; breakout by CSP, ECS)?	How many teachers teach a CS course?	E/V	E/V	E/V	E/V
	How many teachers teach each CS course?	E/V	E/V	E/V	E/V
	How many teachers teach multiple courses? 1. 1, 2, 3, 4+	E/V	E/V	E/V	E/V
	Do you identify teachers who teach CS courses? 1. Gender, ethnic and racial identities; disability status; 2. Training history 3. Certification/credentialing 4. Primary teaching field 5. # years teaching K12 6. # of years teaching CS	E/V	E/V	E/V	E/V
Who are (and are not) the students taking and completing a CS course(All CS; breakout by CSP, ECS)?	What are the demographics of the overall student population? 1. Gender, ethnicity, race, language, SES, free/reduced lunch, disability	E/V	E/V	E/V	E/V

	Who are the students taking Computer Science (by course)? 1. How many students 2. Demographics: Gender, ethnicity, race, language, SES, free/reduced lunch, disability	E/V	E/V	E/V	E/V
	Taking at least 1 computer course in academic year? • What are the demographics of these students? • What is the pass rate/demographics of these students?	E/V	E/V	E/V	E/V
How well do student perform in CS courses (All CS; breakout by CSP, ECS)?	Outcomes 1. Passing rates 2. Grades 3. AP scores 4. Prior/Subsequent course taking	E/V	E/V	E/V	E/V
	For all graduating 12th graders, what % have had 0/1/2/3+ computing courses? 1. What are the demographics of students who have had 0/1/2/3+ computing courses?	E/V	E/V	E/V	E/V
	How many students pass CSP and CS-A courses? 1. What are the demographics of students passing each AP exam?	E/V	E/V	E/V	E/V

The Ease and Value scales are as follows:

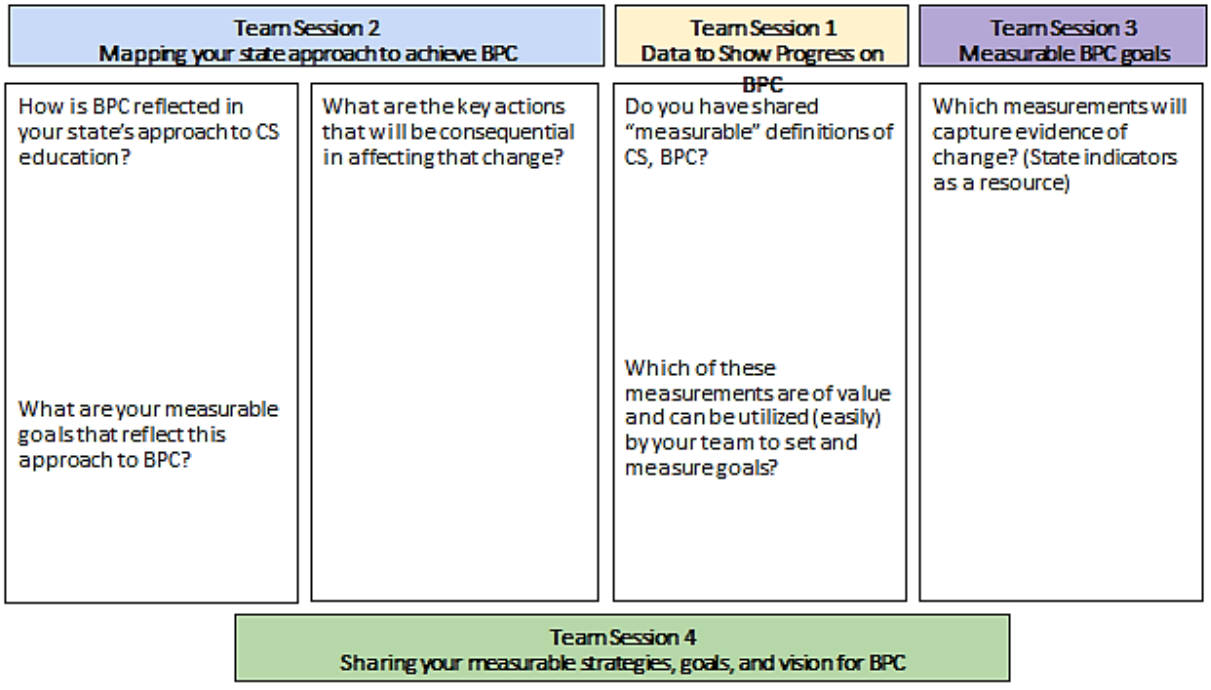
Ease	Value
3 = gathering this information is easy for our state team *example: State has an identified leader who has consistent access to data sources* 2 = gathering this information is somewhat difficult/complex for our state team. *Example: Data is collected based on relationships and not formal pathways*	3 = this information is highly valuable to our state team 2 = this information is somewhat valuable to our state team

1 = gathering this information is very difficult/complex for our state team *Example We know what we need to collect but do not have access*	1 = this information is not valuable to our state team

Facilitated Breakout Sessions at the ECEP 2018 Annual Meeting

Prior to the annual meeting, SageFox reviewed the State Data Collection Template results and developed, in conjunction with the ECEP Annual Summit planning team, a quasi-logic model for state data collection. The templates were completed over the course of four breakout sessions over the two day Summit. Teams were broke out into four rooms with 3-4 state teams each to work independently but also allow for cross-talk between states. Teams were provided a template upon which to take notes for each of the four facilitated sessions, drawing from the two pre-meeting templates:

1. Data to Show progress on BPC
    a. Do you have shared measurable definitions of CS and BPC?
    b. Which of these measurements are of value and can be utilized (easily) by your team to set and measure goals/
2. Measurable BPC goals:
    a. Which measurements will capture evidence of change?
3. Mapping your state approach to achieve BPC
    a. How is BPC reflected in your state's approach to CS education?
    b. What are your measurable goals that reflect this approach to BPC
    c. What are the key actions that will be consequential in affecting change?
4. Sharing your measurable strategies, goals and vision for BPC



Participant responses were captured via observation and notes taken during the breakout sessions.