From Data Bonk to Data Wonk: The Value of Collaborative Exploration of State-Based Data Systems in Support of Equitable Computer Science Education Policy, Programs and Practices

Rebecca Zarch
rzarch@sagefoxgroup.com
SageFox Consulting Group
Amherst, Massachusetts, USA

Sarah T. Dunton
sdunton@mghpcc.org
Massachusetts Green High
Performance Computing Center
Holyoke, Massachusetts, USA

Joshua Childs
joshuachilds@utexas.edu
University of Texas at Austin
Rocquencourt, Texas, USA

ABSTRACT

The K–12 broadening participation in computing (BPC) effort requires access to comprehensive state and national K–12 data from which stronger strategies for systems change can be developed. The Expanding Computing Education Pathways (ECEP) Alliance Common Metrics Project (CMP) engages state teams that include state and local education agencies, researchers, and other BPC advocates addressing K–12 computer science (CS) inequities in access and participation at the systems level. The CMP promotes collaboration and knowledge sharing, with teams reporting how CMP enhances BPC policy, pathways, and practices to improve student access and participation in computing. This experience report shares how the CMP advances data as a key tool for driving BPC strategies in state advocacy and policy efforts.

CCS CONCEPTS

• **Social and professional topics → K-12 education**;

KEYWORDS

Broadening Participation in Computing, K–12 Computer Science Education, Equity, Data

1 INTRODUCTION

"People talk data differently" is perhaps one of the most astute observations about the project from Dana Calfee, STEM & Computer Science Specialist at Indiana Department of Education and one of the participants in the Common Metrics Project (CMP). The CMP is a cohort-based model for state teams to collectively create a shared

language and vision for using data to inform policy and practice across an alliance of states working to broaden participation in computing. In sharing the benefits of the project to the alliance, Dana continued that "the way we represented data for our state in the past, it's just a lot different than how we want to do it today."

This experience report describes the practice of working with 16 states plus the territory of Puerto Rico exploring the data available through their State Education Agency (SEA) focused on high school student participation and access in K–12 computing. The project explored how these data can benefit state teams in advancing their policy and programmatic efforts in support of equitable computer science education (CSEd). When considering early efforts to collect and utilize data in service of broadening participation in computing, there were some clear data 'bonks'. These bonks include a disconnect between the data available and the ways that advocates and researchers needed to utilize the data to fully understand the landscape of CSEd. As the project matured, we shifted away from a data bonk culture that passively accepts data at face value, works in isolated silos, and uses data to perpetuate disconnected problems. We moved towards a data wonk culture of collaboration, storytelling, and critical exploration of disaggregated data. We actively acknowledge and address bias and subjectivity, with a central focus on utilizing data as a catalyst for achieving equitable change. By doing this collaboratively, we are becoming data 'wonks'. This experience report seeks to impress upon others engaged in the BPC movement both how challenging and how important it is to develop data practices and to work collaboratively with a broad based team to ensure data is defined, collected, understood and used in an equity-focused manner.

Over the past 11 years of Expanding Computing Education Pathways (ECEP) Alliance work, the need for data sources that accurately reflect the gains and continued gaps in equitable K–12 CSEd has become paramount in the BPC movement. Several states have data monitoring built into legislated CSEd policy efforts but many state teams rely on independent data collection efforts to set strategy, build broad based support and monitor the implementation of CSEd policies and practices. States are asking similar questions of their data and each other when trying to meet their equity goals, such as "who is taking a CS course? Or more than one? Where are courses being offered? And, what courses do we include when making queries about CS courses, access and participation?" Related to this are the equity and contextual questions such as "who's *not* taking CS?" For example, in some places it would be exciting to see 80 students taking CS, but maybe not as exciting in a school of

1,000 students than a school of 300. And, states are wondering if there are disparities between the demographics of students that are participating in CS and the demographics of the general student population, and if these are closing over time. What if one school has 5 sections of a CS class but the neighboring school has none? What would a state do to support and resource a school that would like to build a CS program? These are common questions within and across states. When we talk about scaling CS as part of the BPC movement, we need this information in context. The nuances in the data can make it hard to understand our collective progress without a shared understanding of the data ecosystem.

The CMP began in 2019 when the 6 New England States came together during a regional Computer Science Teachers Association (CSTA) conference with the query, "how are we, as a region, doing in advancing our BPC goals?" After working together across states to understand local data systems and the regional context, the CMP was launched as a signature project of ECEP. The CMP has the dual purpose of developing state capacity to work with their state data and developing a better understanding of state data ecosystems as ECEP develops national resources to support data-driven equity work. Cohorts of 5–7 states are formed primarily driven by region but also considering readiness to participate (requiring a full team with access to data). By the end of 2024, the 29 states and the territory of Puerto Rico that make up the ECEP Alliance will have gone through the CMP process.

The work of ECEP and the CMP are built on the National Science Foundation's (NSF) efforts to broaden participation in the science fields. In *Investing in America's Future: National Science Foundation Strategic Plan FY 2006–2011* the federal government's growing commitment to broadening participation in the sciences was evident. The plan noted the need for all Americans to not only have opportunities, but to engage with communities underrepresented in the sciences in order to ensure competitiveness in research and workforce [3]. It is from this work that ECEP has scaled efforts across 29 states and the territory of Puerto Rico aimed at increasing the number and diversity of K-12 students in computing. Through both research and implementation across 12+ years, ECEP has begun to focus more closely on the relationship between data and educational policy reform. Over the six years of the CMP, and four cohorts of state teams, the lack of consistent data, and consistently accessible data is evident. State leaders have noted that the lack of access to good data impedes their ability to frame a complete story of the disparities in K-12 CSEd. ECEP and the CMP address this challenge by working to identify barriers to data access, and a lack of capacity within states that makes data utilization a challenge.

The CMP is unique from other data efforts because the core of the work is developed in partnership with the state leaders who will utilize the data to drive change efforts. A key outcome of the CMP is the ability for state teams to build their capacity to utilize data in service of their BPC efforts, as well as publicly share disaggregated data that highlights outcomes around access and participation for prioritized underserved student populations. Despite the long investment in CSEd, the CMP was often the first time a broad based group of state-based stakeholders had to define what CS meant in their state. This level of attention to data empowered state teams to be more confident in their ability to utilize data, and at the same time helped them to see opportunities to improve access to data for strategic planning.

2 DESCRIPTION OF PRACTICE

The CMP requires the participation of highly engaged state teams which include people from State Education Agencies (SEAs) and Local Education Agencies (LEAs) responsible for data gathering, data analysis, data reporting, and data utilization. The CMP brings these teams of 3–7 state representatives committed to CS educational equity from each state through a one-year, four phase effort that covers:

CMP Team formation: CMP teams are often slightly different than the core ECEP team as they require people who can access data through the Departments of Education, interpret the data, and make useful meaning out of this data for policy makers and advocates.

Exploration of the data ecosystem: Teams need to understand how data moves from the classroom to the district and ultimately to state data systems. Mapping the data flow uncovers the variation within and between state systems including: what's possible when seeking to use state data for an equity purpose; the limitations; identifying points of vulnerability; and how to center equity.

Collaborative learning across states: This learning environment promotes trust, builds consensus around "common enough" definitions and shared practices for collecting and interpreting data including subjectivity and bias in reporting. The concept of "common enough" allows state teams to step outside of their often rigid data systems and definitions and promotes a space from which to build and engage [2].

Tying data to practice: Ensuring data is collected, analyzed and utilized in support of BPC requires a commitment to turning data into actionable activities and practices. The orientation elevates models for data utilization and highlights best practices for the democratization of data.

An important part of the CMP philosophy is that these broad-based state teams work together to develop a sense of shared ownership and understanding of the data available to them through the state and make decisions about how this data will be used to tell their CS story in support of equitable policy, programs, and practices. One of the challenges teams confront is that there is very little control of the data input, both from a content perspective (how the system defines and gathers each data element) and a quality perspective (how the individuals selecting course codes, for example, decide which code applies to which class). Yet state level data which each state mandates be gathered is a relatively reliable and sustainable data collection effort making it an important tool for those invested in equitable CSEd. Through CMP, we work to deeply understand its limitations so we can use the data available in new, equity-driven ways. This discussion occurs within state teams, but also across state teams. Unlike other national data collection efforts, the CMP centers state data culture and local context. It leverages the preexisting ECEP Alliance community to develop the data collection templates and definitions that are core to the CMP success. Top down decision making does not restructure the data.

This intentional community engagement ensures that the data is relevant, meaningful, and not manipulated in ways that make the final product unrecognizable to states.

As a learning community, ECEP promotes sharing across the Alliance. At the conclusion of each cohort, the participating states report out how they have applied their experience within the CMP to their state BPC work. An examination of the transcripts of the first three cohorts of states surfaced seven profound ways in which the project has advanced BPC.

2.1 Clear definitions of CS

Despite the long investment in CS education the CMP was often the first time a broad based group of state-based stakeholders had to define what CS meant in their state. When working with data, clear definitions of what counts as high quality CS is a fundamental first step. Jen Rosato, Director of the Northern Lights Collaborative for Computing Education in Minnesota realized "we haven't had to operationalize our definition of CS at all, we haven't had to wrestle with what is and is not CS" prior to this project. Having a definition allows states to organize around a common frame of reference. This definition allows states to assess the values driving the investment in CS for the state, set strategy, create meaningful data queries, and track consistent data over time [6].

2.2 Data quality

One state team member shared that "one of the great things about this project is it forced us to look at the quality of the data." Many team members had never interrogated the quality of the data that was entering the state systems. Points of vulnerability from course code application to local point-of-entry, and working across often disconnected systems were uncovered. States created professional development for their districts to more accurately assign course codes to ensure better consistency and alignment across the states. For example, Rhode Island and Connecticut both created guidance documents and workshops for local administrators to more accurately apply course codes. In 2021 the Connecticut Department of Education released the Computer Science Course Reporting Guidance document that was also accompanied by a video and regional training series [5]. The CS4RI group published "CS4RI — Data Collection Guidance & FAQs", which was updated in 2022 [1].

2.3 Disaggregated data driving better strategy

States began to look at regional needs as data was disaggregated. In many states the overall data told a different story than data that was broken down regionally or by demographic group. Examining data in this way helped states set priorities for interventions. Angela Oechslie, Program Director of Educate Maine noted the Maine team will pursue this further when she shared "we definitely identify the need for a computer science landscape analysis, we think we got funding for that." Emily Thomforde, former Maker Education and Computer Science Coordinator at the San Mateo County Office of Education realized that their state team "really cared about things like low income status, disability status, and [English language learners], which are not necessarily reported by our state" which made it difficult to understand how well their CS strategies were serving specific student populations as they set strategy.

2.4 Data to monitor policy

Some states have enacted CS policies that require monitoring through data. These states are now better equipped to meet this need. Sarah Carter, STEM/Computer Science Specialist at the Minnesota Department of Education realized the Department could "add CS data to an existing report" which was an important efficiency.

2.5 Improving dashboards

States learned from each other how to initiate the development of or improve their dashboards. Being together, collaborating across states, was a significant source of inspiration. Brett Tanaka, Education Specialist at the Hawaii Department of Education noted that "I was really inspired by that first dashboard presented by Georgia...It really set the tone for when the metrics process came along that we were looking for." Establishing a dashboard with an equity focus early is an important and collaborative effort. Many teams are accessing data that is not typically intended for equity work, may have a significant lag before becoming publicly available and/or struggling to create systems that are user-friendly and compelling to diverse audiences [4]. This heavy lift is eased with a community that shares the same equity-based and user-focused principles.

2.6 Developing/refining data-driven strategic plans

A strategic plan supports the collective impact of a broad based team in making systematic change in equitable CSEd. Often teams are comprised of people from state and local education agencies, non-profits, policy advocates and researchers. States developed or refined their strategic plans based on available data, with each case needing different data, or stories, to address unique audiences. Sarah Carter said "we are just at the beginning stages of creating our very first state plan for CS, and to have that data at the front end is really fantastic." Andrew Cronk, Computer Science Education Specialist at the Oregon Department of Education reported "...looking at data together, understanding the nuances of the experiences across the state, looking at intersectional data, and looking at the experiences of our non-binary data truly helped us center equity in our plan. And now, I'm very happy to report we've just released the first draft of our state plan for public feedback."

2.7 Improved messaging through data

The CMP emphasizes data as a tool for supporting equitable policy, practice, and strategy, not just to report the current status. Using data to inform a wide range of advocates within a state is important, and often requires differentiated messaging based on the perspective of the individual, however having a centralized data system (dashboard or landscape report) can streamline this process. Helen Hu, Processor of Computer Science at Westminster University in Utah shared that they "often meet with legislators and other district leaders" and knowing their needs helped define the dashboard creation, but importantly "one of the reasons people were willing to work with us is we kept talking to them about how our goal is to create one consistent data request" that can be used each year for sustainability. Another state realized "you can talk about data till the cows come home, but unless it's really conveyed in a way

that's going to stick with people and stay with people, that's the way differences are going to be made."

3 POSITIONALITY STATEMENT

This paper was authored by two white women and a Black man who serve on the ECEP Alliance senior leadership team. We recognize that race, gender, socioeconomic status, sexual orientation, religious background, professional and educational experience, and our organizational affiliations are just a few identity markers that shape our lived experiences. We recognize the necessity of being transparent about our biases, and the urgency of continuously learning, growing, and listening to our colleagues, peers, leaders in the field, students, educators, and those with different experiences than ours. We see our role within ECEP and CMP as 'facilitators', where we amplify the expertise and experience of state team members. This paper serves as an example of CMP facilitation, where state team members, through a thoughtful designed learning process, rely on their collective expertise and experience to support new ways of collecting, sharing, and supporting equity-focused CSEd data. In addition to experience as researchers, our facilitation is guided by our experience as an evaluator focused on equity in STEM education (Zarch), the Director of the ECEP Alliance (Dunton) and a Professor of Educational Policy and Planning (Childs). We are also people who have career paths that include coaching school age children, leading youth development programs, and working in educational settings serving youth of all ages.We acknowledge that none of the authors currently work directly within the K–12 system. Therefore, how we conceptualize and use data does not put our employment at risk, yet we are still impacted by destructive Diversity, Equity, and Inclusion (DEI) laws and policies that can influence how we interact with state team members, our daily lived experiences, and professional interactions that may restrict how we approach important conversations around equity-oriented justice within CSEd.

4 LIMITATIONS AND ASSUMPTIONS

This project privileges state based data, typically collected by state DOE offices, that focuses on public K–12 schools. State-based data systems and dashboards are inherently limited and may even create barriers to educational equity. Some state-based data systems may only highlight a part of the data story about a community, school, or the state. Too often the use of data (from how it is captured to how it is used) focuses on student deficits or serves to 'other' differences and disparities. The CMP model is designed to democratize the process of gathering and using data in ways that allow state teams to identify harmful data practices, collect and analyze other types of data to tell a more complete story, and set future strategies to build partnerships that can influence improved data collection and analyses. These additional strategies recognize the importance of having robust data that intersect with other social sectors such as transportation, housing, and public health indicators. Furthermore, the CMP model highlights the importance of localized knowledge, where leaders directly working within and with communities (i.e. students, educators, etc.) can acknowledge and capture data that reflects that experience and capacity of a community as it relates to broadening participation in computing. The CMP project does not define how states move forward, recognizing that this level of decision making should be locally driven and locally dependent.

The data shared in this report comes from fairly open-ended community report-outs and not a systematic research study of the impact of the program. Had we used a more formal research protocol, states may have chosen to speak about their experience differently, perhaps emphasizing different elements of the experience.

5 IMPLICATIONS AND NEXT STEPS

Developing a data wonk culture, that includes data literacy, data utilization in service of equity, and awareness of positionality, requires trust and collaboration. As Dana Calfee found, people do talk data differently. Creating a shared language doesn't just mean using common words but having a common understanding of the data ecosystem and how data can be used to make change. Creating and using data for policy change depends on a broad based team. Since we began the CMP project with the New England states in the fall of 2019, several teams have seen significant turnover at the leadership level. Teams are constantly evolving and data work is never complete. A robust data culture requires continuous investment and support as new directions can be set and new individuals are brought to the table. As the CMP moves forward we have identified three priorities for the ECEP Alliance data efforts. The first is to gather the learning across the states into a revised data landscape guide to support data teams widely. The second is to explore new modes of data visualization to advance storytelling and the third is to explore an Alliance-wide data dashboard that will further democratize state level data. Each of these approaches will support and empower our states as they build collaborative data cultures in support of equitable CSEd policy. We plan to develop "ECEP Briefs", short research-based publications, to help share concrete examples and practices from state teams that may assist other teams in their BPC efforts.

Data is a type of storytelling. One way to make data stories more compelling is to visually represent the data in a variety of ways for BPC impact. The CMP provides the data infrastructure that allows teams to tailor the data to a wide range of audiences as needed. It also engages teams in uncovering new questions or avenues for exploration. For example, in our most recent round of CMP interviews we universally heard the need for understanding the teaching capacity within a state. As we move forward in our own project we are exploring new methods for engaging people through data visualization including interactive dashboards and infographics.

Working across states is similar to other projects that try to aggregate data sets across disparate sources. Data that has not been structured or defined in common creates inconsistencies leading to difficult interpretation and interrogation [7]. It also raises questions of data privacy which is complicated by each state having their own interpretation of the Family Educational Rights and Privacy Act (FERPA). Given the progress made through the CMP, however, we are confident that creating an Alliance-wide dashboard is a feasible and necessary step for centering data in the BPC and systems change mission of ECEP and its alliance of state teams.
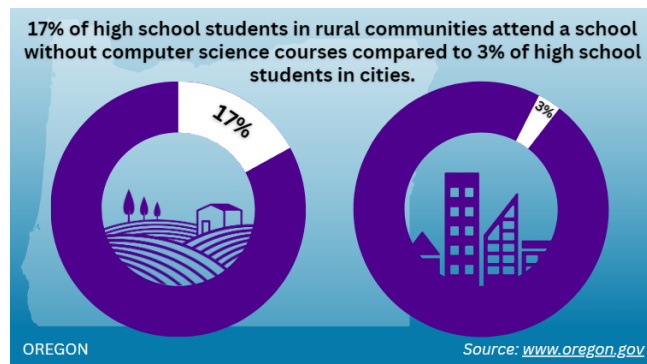
Figure 1: An example of data visualization used in the CMP

ACKNOWLEDGMENTS

REFERENCES

[1] CS4RI. 2022. Data Collection Guidance & FAQs. https://www.cs4ri.org/data-guidance

[2] Sarah Dunton, Rebecca Zarch, Jeffrey Xavier, Jayce Warner, and Alan Peterfreund. 2022. Determining metrics for broadening participation in computing: Connecting data to multi-state computer science education policy efforts. *Policy Futures in Education* 0, 0 (2022). https://doi.org/10.1177/14782103211064443

[3] National Science Foundation. 2006. *Investing in America's Future: National Science Foundation Strategic Plan FY 2006–2011.* Technical Report. https://www.nsf.gov/pubs/2006/nsf0648/nsf0648.jsp

[4] National Academies of Sciences, Engineering, and Medicine. 2019. *Monitoring Educational Equity.* The National Academies Press, Washington, DC. https://doi.org/10.17226/25389

[5] Connecticut State Department of Education. 2021. *Computer Science Course Reporting Guidance.* Technical Report. https://portal.ct.gov/-/media/sde/computer-science/course-code-guide-final.pdf

[6] Rebecca Zarch and Sarah Dunton. 2022. Looking Back to Move Forward: Measuring K–12 Computer Science Education Requires an Equity-Explicit Perspective. In *2022 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT).* IEEE, 100–104.

[7] Rebecca Zarch, Alan Peterfreund, Burçin Campbell, Heather Wright, Monica M. McGill, Julie M. Smith, Talia Goldwasser, and Tyler Clark. 2024. *Building Shared Measures for Broadening Participation Initiatives.* Technical Report. https://doi.org/10.62572/jZamU35