Evaluating Computer Science Professional Development for Teachers in the United States

Monica M. McGill
CSEdResearch.org & Knox College
Peoria, IL, USA
monica@csedresearch.org

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

Stacey Sexton
SageFox Consulting Group
Amherst, MA, USA
ssexton@sagefoxgroup.com

Julie M. Smith
SageFox Consulting Group
Amherst, MA, USA
jsmith@sagefoxgroup.com

Christine Ong
UCLA CRESST
Los Angeles, CA, USA
ong@cresst.org

Melissa Rasberry
American Institutes for Research
Durham, NC, USA
marasber@ncsu.edu

Shelley Hollis
Mississippi State University
Mississippi State, MS, USA
shelly.hollis@rcu.msstate.edu

ABSTRACT

Teacher professional development (PD) is a key factor in enabling teachers to develop mindsets and skills that positively impact students. It is also a key step in building capacity for computer science (CS) education in K-12 schools. Successful CS PD meets primary learning goals and enable teachers to grow their self-efficacy, asset and equity mindset, and interest in teaching CS. As part of a larger study, we conducted a secondary analysis of CS PD evaluation instruments ($n = 14$). We found that instruments across providers were highly dissimilar with limited data collected for measures related to teacher learning, which has implications for future K-12 CS education. Likewise, the instruments were limited in being connected to student learning and academic growth. As a way to enable PD providers to construct measures that align with known impacting factors, we offer recommendations for collecting demographic data and measuring program satisfaction, content knowledge, pedagogical content knowledge, growth and equity mindset, and self-efficacy. We also highlight questions for PD providers to consider when constructing their evaluation, including reflecting community values, the goals of the PD, and how the data collected will be used to continually improve CS programs.

CCS CONCEPTS

• **Social and professional topics** → **Computing education**; **Computing education programs**; **Computer science education**.

Koli Calling '21, November 18–21, 2021, Joensuu, Finland

ACM ISBN 978-1-4503-8488-9/21/11.
https://doi.org/10.1145/3488042.3488054

KEYWORDS

Professional development, evaluation, teacher, K-12, primary, secondary, recommendations, self-efficacy, mindset, content knowledge, pedagogical content knowledge, PCK

ACM Reference Format:
Monica M. McGill, Rebecca Zarch, Stacey Sexton, Julie M. Smith, Christine Ong, Melissa Rasberry, and Shelley Hollis. 2021. Evaluating Computer Science Professional Development for Teachers in the United States. In *21st Koli Calling International Conference on Computing Education Research (Koli Calling '21), November 18–21, 2021, Joensuu, Finland*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3488042.3488054

1 INTRODUCTION

Professional development (PD) is an integral part of training K-12 teachers and has been shown to positively impact teachers' content knowledge (CK), pedagogical content knowledge (PCK), self-efficacy, and beliefs [24, 45]. This in turn impacts *students'* CK, self-efficacy, and beliefs [24, 27, 59]. PD is, in fact, vitally important when it comes to expanding Computer Science (CS) education in K-12 schools.

In the U.S., schools across 50 states and 13,000 independent school districts continue building capacity to collectively teach computing to over 50 million K-12 students [41, 58]. The fact that the U.S. does not have a national curriculum complicates discussions of how computing should be taught and what content should be delivered to students. It also complicates efforts to train teachers across districts and states. Further, there are over 150 unique providers offering various types of CS PD across various grades [14]. Deciding what CS PD to offer, how to offer it, and who to offer it to can make delivering and evaluating CS PD challenging [36]. These challenges raise two questions:

- *What data do CS PD providers collect to evaluate their CS PD?*
- *How do these compare to factors that are known to impact student learning?*

We conducted our research on evaluating teacher PD as part of a larger study examining how CS PD providers in the U.S. pivoted

from in-person to virtual during the COVID-19 pandemic. In this paper, we focus on how evaluation of teacher CS PD aligns with factors that impact student academic and personal growth, including standard measures for demographic, program satisfaction, and construct data (e.g., CK, self-efficacy). Based on our findings, the literature and the theoretical model known as the Novice to Expert Theory [9, 19], we offer a set of recommended standards for evaluating teacher CS PD. Our recommendations may also be useful for CS PD providers in other countries with disparate, regionally-based curriculum, which is the current, disjointed model of CS curriculum in the U.S.

2 BACKGROUND

We highlight here teacher factors that impact student learning and academic growth in CS as well as CS PD evaluation. We also provide framing of our underlying theory.

2.1 Theoretical Framework: Novice to Expert Model

Benner and Dreyfus converted a well-known theory about knowledge development in the nursing field to be used in the field of education [9, 19]. Their specific theory covers the trajectory of teachers' movement from novice to expert in their field (i.e., novice, advance beginner, competent, proficient, and expert). This theory states that professionals move through five stages of career development, which has an impact on future success and career sustainability.

2.2 Teacher Impact on Students

It is well known that teachers influence student learning more than any other school-related factor [16, 28]. The teacher factor has an effect size two to three times greater than any other factor measured within a school environment (e.g., leadership and services) [42]. Among attributes that teachers can develop that positively impact student academic achievement include being committed to and persistent about their students' learning, developing their self-efficacy, and contributing to their professional support system or professional learning network (PLN) (Table 1) [35]. Many of these impacts have been examined with respect to general education and specific subjects (e.g., math, literacy) [31, 49]. From these data there is some empirical evidence to date on their direct impact on students' learning CS [45, 67].

Most notable of these factors are self-efficacy, CK, and PCK. Self-efficacy has been recognized as a crucial factor in pre-service teachers learning CS [60, 64]. Likewise, there is some early empirical data demonstrating that CK and PCK are linked to student academic outcomes in CS [24, 45]. Beyond cognitive factors (e.g. grades and CK), there are many noncognitive factors that teachers can impact, including student motivation and student self-efficacy (see Table 2).

2.3 Evaluating Teacher PD

CS PD providers recognize the impact that their CS PD can have on teachers and their students. Evaluating CS PD is a critical part in developing high-quality, equitable CS PD [4, 17, 23, 36, 39, 66]. Evaluating CS PD, however, remains a persistent challenge [39].

Table 1: Teacher factors known to influence student learning. * indicates factors known to influence students in the context of CS.

Teacher factors impacting student learning
Commitment to and persistent about student learning [61]
Confidence* [24, 25, 29]
Content knowledge* [8, 24, 30, 45]
Content-specific feedback strategies [5]
Feeling good about their professional support system [50]
Feeling good about their teaching [50]
High drive [25]
Interest in teaching the subject* [29]
Pedagogical content knowledge [8, 13, 34]
Perceived collective efficacy [6, 22, 61]
Positive relationships with students [38]
Self-efficacy [31, 49, 53, 65]
Teacher affiliation [26, 46]
Teacher affiliation with their school (belongingness) [26]
Teacher empowerment [52]
Teaching experience [11]
Technological Pedagogical Content Knowledge [1]
Trust, collegiality, and closeness among teachers [55]

Table 2: Teacher factors that have been found to influence student growth (noncognitive factors).

Teacher factors impacting student growth
Attitudes -> Students' Personality [56]
Classroom Organization -> Student Behavior [12]
Communication -> Student Behavior [21]
Encouragement -> Student Self-efficacy [54]
Enthusiasm -> Student Staying on Task [10]
Self-efficacy -> Student Motivation [40]
Pedagogical content knowledge -> Student Motivation [30]
Personality -> Student Self-efficacy [32]

The Computer Science Teachers Association (CSTA) uses a rubric for CS PD providers who want to post their offerings on the CSTA website [14]. The purpose of this rubric is not to formally evaluate program success, but rather to define a minimum level of quality that CS PD offerings should meet, including a content focus, a pedagogical focus, and ongoing support for teachers–all of which correlate to student learning.

How, then, should providers measure their CS PD? And what exactly should they measure, particularly given that there are different purposes for measurement or assessment? Surveys are the essential form used to evaluate teacher CS PD, followed by assessment of CK, and by interviews [37]. Content wise, McGill et al. suggest evaluation across four key areas: CK, PCK, efficacy/beliefs, and program evaluation. This mirrors Merchie et al. [39] and Banilower et al. [7] who recommend similar groupings.

Knowledge, skills and beliefs are common constructs measured in teacher CS PD [3, 39, 43], which generally map back to factors that

impact student learning and academic growth. These constructs can be measured through deliberate and focused CS PD evaluations, as outlined byMerchie et al.. Through their meta-analysis, an extensive effectiveness framework for measuring CS PD was designed. The framework included three critical areas for CS PD evaluation: 1) cognitive goals (teachers' CK and PCK), 2) skills ("what participants are able to do with what they have learnt") and 3) affective goals ("attitudes or beliefs about teaching and learning and beliefs about themselves") [39, p. 8].

3 METHODOLOGY

This study was part of a broader study investigating changes CS PD providers made when pivoting from in-person CS PD in 2019 to virtual CS PD in 2020 due to the COVID-19 pandemic. We collected demographic, outcome and reflective data from evaluators and administrators of CS PD programs using a CS PD inventory that we jointly developed in late 2020. Examples of the inventory questions are provided in tables throughout this report. Providers submitted their responses through a survey platform and were asked to draw from demographic and evaluation data, agendas, and their own reflections. The inventory asked for a description of CS PD activities, including changes made, the process for transitioning to online CS PD, and lessons learned. In addition to a number of different items collected, we asked for specific data related to outcomes as well as the methods and instruments (e.g., surveys, assessments, interview protocols) they used to collect their data.

Utilizing [9] and [19]'s theory, we were able to expertly evaluate CS PD and provide recommendations based on the trajectory of a teacher moving through this model of career development and CK [9, 19]. Utilizing their theoretical framework, we considered critical questions focused on teacher growth in their field (e.g., CS CK and CS PCK). This included questions such as: What recommendations will guide teachers movement up the Novice to Expert Model trajectory? For quantitative data, what instrumentation will provide data on teachers' level of CS CK and PCK?

Evaluating CS PD offerings (e.g., measuring its merit) is different from gathering information that allows providers to understand the CS teachers' current level of knowledge based on learning goals. However, these concepts do overlap. For the purposes of this study, we are primarily focused on evaluation of the CS PD offering; however, our recommendations may or may not apply to the latter.

3.1 Participant Recruitment

Recruitment began in late 2020 after ethics board (Institutional Review Board) approval. We extended invitations to professional networks (e.g., newsletters, announcement at partner meetings, social media). Providers interested in participating completed an online form and we contacted respondents to verify eligibility. To have been eligible, the CS PD provider needed to meet the following requirements:

- Prepared teachers to implement CS curriculum in a K-12 classroom
- Offered in 2019 in person
- Offered CS PD in 2020 that included some synchronous virtual activities
- Was at least approximately one work day
- Changed the instructional model from face-to-face to virtual and/or curriculum in response to COVID-19

Once eligibility was established, we asked providers to complete the 300+ item inventory, offering stipends for participants' time curating and entering data. Five types of CS PD providers submitted data for this study (Table 3).

Table 3: Model of CS PD offered with number of providers and participants.

Model of PD	# (%)	# Part.
Locally-based offering a locally-developed PD	8 (28%)	391
Locally-based PD as part of larger PD effort (e.g., Bootstrap, Code.org, ECS)	7 (24%)	230
Multiple PD opportunities as part of a larger PD effort (e.g., Code.org's CSP and CSD)	5 (17%)	254
Multiple PD opportunities as a mix of locally-developed and as part of a larger PD effort	1 (3%)	865
National providers (summative inventory of all sites)	8 (28%)	4,066

The inventory asked providers to submit the instruments and/or protocols used to evaluate their virtual CS PD. We received 14 instruments for review from the 29 providers. We conducted a secondary content analysis focused on the demographic data items, program satisfaction items, and other constructs they were designed to measure.

4 RESULTS

We asked each provider if they collected seven common constructs for teacher CS PD: Program Satisfaction, Self-Efficacy, CK (self-reported and assessed), PCK (Self-reported and Assessed), and Interest in Teaching CS. Program Satisfaction is important in ascertaining how the CS PD participants perceived the PD experience and how it can be improved. The remaining six constructs were selected due to the prominent role each of them played in the success of CS teacher PD and in student academic achievement (see Section 2). Nine of the 29 CS PD providers (31%) did not collect any data related to these constructs, while 16 (55%) collected data related to program satisfaction. Self-efficacy data was only collected by 8 of the 29 (28%) and interest in teaching CS was only collected among 5 (17%) of these providers.

There are a variety of protocols that can be used to measure CS PD, such as a pre- and post-survey, post-survey only, intermittent surveys, longitudinal surveys, etc. Reliability and validity are measures for ascertaining the quality and accuracy of an instrument. We learned that 7 (44%) of the 16 respondents who collected data (in general, whether or not they provided us with related data) used instruments that had prior evidence of validity and reliability.

4.1 Data Collected

Table 4 shows the broad range of demographic data collected, falling into four categories: general demographic data, professional attributes of the teachers, attributes of the schools where the teachers, and attributes of their students.

Ten items related to program satisfaction were measured across the 14 instruments focused on program satisfaction (Table 5). The measurements included questions related to logistics, satisfaction with facilitator/instructor, ways to improve the CS PD, and more. The instruments revealed a wide variety of data collected. This is expected, given the various goals that local, regional, and national CS PD providers may have, as well as the alignment that evaluation constructs may have with those goals. However, over 2/3rds of those completing the inventory rated "Tools to evaluate online CS PD (e.g., survey instruments, data collection strategies)" as medium or high priority for which they need support, we present summary information in Table 6.

5 DISCUSSION

Merchie et al. discuss the importance of triangulation of data when evaluating CS PDs. This includes a mix of quantitative and qualitative data, including observation and interviews.

5.1 Demographic Data

Although there were quite a broad range of demographic data collected, we learned that providers did not collect a number of items that previous research studies have found to be important, including attributes of the schools were teachers taught. Some providers may collect this information as part of their registration or other forms sent to participants. However, though mapping this back to the survey instrument is needed if analysis will be conducted across various demographics. With the continued and persistent call for CS for all students, it seems important for providers to understand if teachers are teaching students with disabilities or if they teach historically marginalized students [39]. By collecting these attributes, it provides context for the teacher and the school in which the teacher is affiliated. It can also give important feedback to providers on how well they are meeting the needs of teachers with varying student needs.

5.2 Program Satisfaction

Program satisfaction includes what Merchie et al. refers to as structural features of the CS PD. These include elements such as duration, trainers/facilitators, pedagogy, and collaborative participation. Closely related, Doppelt et al. identified their own set of features that impact teacher learning, including CS PD sessions that are distributed across several weeks, "...engaging teachers in an active learning process situated in the curriculum; and facilitating a collaborative community of teacher professionals." [18]

5.3 CK, PCK, and Beliefs Data

If we compare teacher factors that impact student learning and academic growth (Tables 1 and 2) with constructs measured (Table 7), we find a significant overlap. Collectively, the instruments measured these related factors. We note that in both CK and PCK, the instruments relied on self-reported measures of growth, rather than actual assessed growth.

Merchie et al. considered a framework for measuring teacher CS PD and have more frequently found in their systematic literature review that CK is typically measured through content tests and questionnaires rather than self-reported. Self-reported assessments are often inaccurate, and especially inaccurate for different populations (e.g., women having less confidence in their knowledge) [33, 47].

5.4 Limitations

Since we extended our invitation to those who offered CS PD in the summer of 2019 and 2020 and who met other criteria (see Section 2), the set of instruments we evaluated was limited. A less restrictive call for instrumentation could further provide information on what type of data is typically collected in teacher CS PD. We recognize that evaluating CS PD can occur informally, with informal feedback from facilitators and instructors observing the behaviors and engagement of teachers during the CS PD.

We also interpreted what the items were intended to measure based on the items themselves, rather than asking providers to explain what constructs certain items were set to measure. Thus, the intent of the items may differ from our interpretation. Some of this required a professional judgment call for which we relied on our experiences and expertise with instrumentation, survey questions, and constructs.

We did not measure the number of instruments that collected each construct. The levels of measurement within the instruments differed, and a count of these would not accurately represent the strength of each measure. For example, one instrument may have only asked one single item related to self-efficacy, while another may have had a set of six questions formed from an existing self-efficacy instrument. However, very few instruments measured more than three constructs and (self-reported) CK was commonly asked.

6 RECOMMENDATIONS

As is evident in the literature and the data we gathered, CS PD providers should be clear about assessing and evaluating teacher experiences and knowledge in CS. This ensures that CS PD is purely focused on improvement and not evaluating teacher knowledge. Improvements should based on the Novice to Expert Model [9, 19]. Further, understanding the difference between evaluation and assessment is essential, since the words *assess* and *evaluate* can have different meanings, some of which can impede teachers' sense of identity, worth, and sense of belonging. This can unintentionally impact teachers' honesty in completing any survey or form.

There have been various efforts to hold teachers accountable for their students' learning (i.e., use students' scores on various standardized tests as a measure of their effectiveness). Teachers have been found to be more adversely impacted by any form of assessment or evaluation that appears to be assessing skills. Due to this phenomenon, CS PD offerings occurring within a school district may *feel* threatening. We took this into consideration when developing the recommendations for the types of constructs and approaches used to evaluate CS PD.

Table 4: Types of demographic data collected from participants. An * indicates data we recommend providers collect as a standard part of data collection. Other attributes can be collected on an as-needed basis.

Teacher General Demographic Data	Teacher Professional Attributes	School Attributes	Student Attributes
Gender*	Current role*	*Access to Lab for class	*Experiences with CS
Race/ethnicity*	Years taught*	*Only CS teacher at school	Number taught per year:
Disability*	Years taught CS*	*Permission to have software installed on school computers	*All Students
	Volunteered or Required to teach CS*	*Rural/urban/suburban school	*Girls & nonbinary students
	Access to PD	*Socio-economic status of school	*Historically marginalized students
	Certified to teach CS	CS courses offered at school	*Students with disabilities
	Credential area	Prominence of CS (elective, integrated, required)	
	CS courses taught	School name/location	
	Degree Subject Area	Student recruitment process	
	Grades certified to teach		
	Grade levels taught		
	Plans to teach CS		
	Previous CS PD		
	Previous CS training		
	Subjects taught		
	Years in Education		

Table 5: Measurements of Program Satisfaction

Area	Count (%)
Logistics (e.g., registration, communication)	9 (64%)
Satisfaction with facilitator/instructor	5 (36%)
How they learned about the PD	3 (21%)
Most rewarding aspects of the PD	3 (21%)
Sense of community/belongingness	3 (21%)
Ways to improve PD	3 (21%)
Reasons for attending the PD	2 (14%)
Engagement in the course	2 (14%)
Pace of the course	2 (14%)
Most challenging aspects of the PD	2 (14%)
Stability of online environment	1 (7%)

We also recommend determining how much time is focused on constructive CS PD activities, such as looking at student work, talking about assessing student understanding, and having teachers reflect on their own learning. When deciphering how time is spent, it is important to reflect on: What are the goals of the each teacher coming into the CS PD? How do the unique nuances of teaching computing require teachers to think and teach differently?

Measuring CS PD against a particular program's goals requires reflection of the providers' values [36]. McGill et al. provide key questions to consider when deciding what to measure:

- What are the values of the district and school communities in which participants belong?
- What are factors that impact student learning that are most important to the participants?
- What key data points are needed to improve CS PD offerings?
- What constructs around equity should be measured to better understand if the CS PD affects teachers' beliefs about each student's ability to learn CS?
- From which other constructs that might impact student learning should data be collected?

Not all providers will have the capacity to collect and analyze a broad set of data. Though these recommendations are broad, each provider will need to weigh their capacity for evaluation and prioritize their needs with their resources.

Finally, each recommendation is based on the data analysis of submitted documents, as previously outlined, as well as scholarship regarding CS and PD.

6.1 Demographic Data

Ultimately, CS PD providers must decide which data to collect. Based on our experience in the field and with CS PD, we encourage CS PD providers to collect meaningful data that is inclusive (e.g., dis/abilities of participants and their unique needs) and will lead to an improved experience for teachers by directly meeting their needs. Most of these recommendations are also present in undergraduate teacher education programs as a way to evaluate pre-service teachers movement from novice to expert. Using the breakdown as shown in Table 4, we recommend that the following measures be collected as a *standard practice* for CS PD providers:

Participant General Demography. We recommend that CS PD providers collect participants' gender, race/ethnicity, and dis/ability. This helps gather adequate data for ensuring equitable access to the CS PD. It also gives CS PD providers (as well as others) information about which groups are underrepresented in the CS PD. Though dis/ability is often overlooked, we call this out as necessary data to ensure the CS PD meets participants' needs.

Table 6: Measurements of Teacher Outcomes. Measurement types are grouped based on past recommended categories [7, 36, 39]. * indicates recommended constructs for CS PD providers to evaluate.

Measurement Type	Specific constructs/knowledge measured
Content Knowledge	Computer Science* (self-reported) Computer Science Standards (self-reported)
Pedagogical Content Knowledge	Computer Science* (self-reported)
Beliefs and Efficacy	Ability to manage stress/pressure at school Belongingness Benefits of students learning CS/Purpose of CS education Comfort teaching CS Confidence teaching CS Confidence teaching specific units in CS course Confidence teaching topics/units Confidence with resources that focus on equity Feelings about their school's plan to teach CS Understanding of how biases impact teaching Perception of student engagement in CS course Persistence Self-efficacy teaching CS* Supported teaching CS*
Beliefs related to Mindset	Interest in teaching CS* Equity Mindset* Growth Mindset*
Beliefs related to Teaching Practice	Adaptability to new teaching practices Challenges teaching CS Comfort level with technology/CS Concerns about teaching CS Familiarity with Instructional Coaching How much teacher likes teaching with packaged curriculum Preparedness - Feel prepared to use material learned

Participant Professional Attributes. We recommend CS PD providers collect participants' current role, years taught, and years teaching CS as important aspects of framing the experience. Asking participants whether they volunteered or were required to teach CS can provide insight into a participant's perspective of their PD experience. The other attributes appearing in Table 4 may be valuable to add based on the focus or objectives of the CS PD. Professional attributes are known as dispositions in teacher education programs.

School Attributes. We recommend collecting data about school attributes such as rural, urban, or suburban. This information provides an understanding access to CS PD and whether or not the participant is the sole CS teacher in their school. If the participant is the sole CS teacher in their school or district, this provides context for understanding issues related to community, resources, and support for the participant. Adding to this data, we recommend collecting data about participants' access to a teaching lab or other questions related to their teaching environment, as well as their school's ability to install software on computers for teaching CS. Answers to these questions may affect the software the schools/participants choose to use to teach CS.

Student Attributes. Asking participants about student attributes can be instrumental in understanding the composition of the participants' classrooms and unique challenges participants may face in creating an equitable environment for learning CS. Asking teachers to estimate their students' proficiency or previous experiences with CS can help PD providers understand what pedagogical content to include.

Asking a teacher to *estimate* their students' demographics regarding gender (including non-binary or gender fluid), dis/ability, race/ethnicity, and socioeconomic status presents its own unique challenges. We recommend collecting the *actual* demographic data for the students taught each year.

6.2 Program Satisfaction

What CS PD providers collect with respect to program satisfaction varies across providers. It seems reasonable that providers want to collect information about their logistics (e.g., registration, communications, etc.), and we found that approximately two-thirds of providers do. When program satisfaction data is gathered, it can provide a wider lens of growth or non-growth on the trajectory of CS teachers novice to expert journey. Here, we offer a set

of recommendations for collecting information related to CS PD programming satisfaction as a standard practice:

Logistics. We recommend that providers consider logistics related to processes (e.g., recruiting, communications to participants, ease in registering).

Learning environments. Environments can positively or negatively impact learning. When offering PD virtually, it can be expected that outside pressures may influence learning (e.g., caretaking responsibilities, Internet access, hardware needed for the course). We recommend that several sets of questions be created to ascertain participants' learning environments.

Course Engagement. Engagement is an important aspect of learning, and there may be other measures that can provide clues (e.g., observation, sidebar conversations) as to the level of engagement by participants. However, asking participants what their perceived level of engagement was during the course is also acceptable. We recommend one or two questions related to course engagement be added.

Pace of the course. In-person and virtual course offerings have their own pace, which is an important aspect of learning. We recommend asking participants for their perceptions of the course pace. This will provide valuable data for future CS PD implementation.

6.3 CK, PCK, and Beliefs Data

CK, PCK, and teacher mindset (both asset-based and equity-focused) are key pillars for CS PD providers to target if all students are to learn CS. Without understanding mindsets of the participants, the CK and PCK aspects of CS PD will be inadequate in meeting the needs of all learners. If a CS teacher cannot meet the needs of all learners, their trajectory on the novice to expert continuum is stagnant. Therefore, we recommend that CS PD providers consider collecting data to enable process improvement of their CS PDs, gathering the following data as a standard practice:

Content Knowledge. We recommend that CS PD providers measure self-reported or assessed CK. However we recognize that assessing is usually more accurate than self-reported. Academic growth can be important, so providers will need to ascertain if it is best to collect the data both pre- and post-PD, or only post-PD.

If collecting only post-PD, and using self-reported measures, CS PD providers can measure growth (with an appropriate 5- to 7-point scale) with questions similar to: How much did you know about *topic* prior to the workshop? and How much do you know now about *topic*? Posing both questions can help gauge whether participants are learning about the content topic to the degree that providers had planned.

Self-reported measures of CK can be influenced by the Dunning-Kruger effect [20], resulting in teachers overstating what they know in a pre-survey. This can result in post-survey data falsely indicating learning losses. A CK assessment consisting of several key content questions, perhaps one for each topic, can be created and used. This should not be over-burdensome or too difficult, since that can negatively impact teacher self-efficacy.

Pedagogical Content Knowledge. PCK is as important as CK [51], and we recommend measuring PCK similarly. There are several CS PCK instruments available to assess PCK, with some of those using vignettes [44, 62, 63]. We recommend considering what aspects of pedagogy are most important to the CS PD and ways to measure those, perhaps as self-reported and post-PD only, particularly if the measures that will be collected are quantitative in nature. For collecting self-reported PCK, we recommend post-PD questions asking teachers if they understand the key aspects of pedagogy that the CS PD offering targeted (e.g., pair programming, teaching computational thinking through storytelling, etc.).

Self-Efficacy. Since self-efficacy has been shown to impact student learning [53, 55, 60], it is important to understand the self-efficacy of teachers who have completed CS PD. Ideally, this can be measured pre- and post- intervention to gauge what type (positive or negative) of and how much impact the CS PD had on teachers.

Mindset. Two mindsets can be evaluated: asset-based and equity-focused. An asset-based mindset encourages teachers to focus on students' strengths and build students' knowledge based on those strengths. Equity mindset "...refers to the perspective or mode of thinking exhibited by practitioners who call attention to patterns of inequity in student outcomes. These practitioners are willing to take personal and institutional responsibility for the success of their students, and critically reassess their own practices." [57] Both mindsets share similar qualities and require practitioners to understand that each student has the capacity to learn, regardless of their personal characteristics, upbringing, family's economic status, and more.

Interest in teaching CS. Many teachers are told that they will be teaching CS, regardless of previous indications of interest. Teachers' interest in the topic being taught can impact student learning [29]. It may be helpful to know whether CS PD offerings can successfully enable and promote interest among CS PD participants, particularly those who have not chosen to teach CS. Inquiring about teacher interest in CS and/or choice to teach CS will be beneficial to providers in meeting the needs of the CS PD participants.

CS Relevance Benefits of students learning CS. Teachers' understanding of how CS is relevant to students and the important benefits of students learning CS can increase their interest in teaching CS. Gathering information regarding teachers' understanding of CS connections to students' realities is another recommendation.

Support network. Isolated teachers (e.g., being the only CS teacher in a school) can negatively impact self-efficacy. Teachers' support networks have a relationship to academic achievement [26, 55]. Being supported promotes a sense of belonging, which is important for teacher self-efficacy.

6.4 Other Considerations

The need to collect data for evaluating and improving CS PD is dependent on available resources and adequate instrumentation. For CK and PCK, the constructs may be somewhat ill-defined in CS (as they are in mathematics [48]) and difficult to measure without a full-blown assessment [2], which we do not recommend. It is important for providers to recognize that CS PD will and should cause some disequilibrium for participants, since change often feels risky. However, it is important to promote and create a safe learning space [15]. If teachers feel that they are being personally evaluated on their knowledge of a topic (or not fully understanding how that information may be used), they might feel less safe or willing to take chances while engaging in the PD.

Other questions to consider when developing an evaluation plan include:

- What data should be collected pre- and post-intervention versus post-intervention only?
- How will CS PD providers ensure that even the least engaged in their CS PD offering provide data?
- For surveys, how long will the survey be? What trade-offs are there in time to take and complete the survey versus amount and type of data collected?
- How can participants be reassured that their data will be used to inform CS PD rather than to evaluate them or their teaching?

Some providers might not collect data, since they then may feel responsible for being able to make change based on that data, yet do not have the capacity to make those changes. For instance, with respect to teacher isolation, a provider might not be able to meet the needs of a teacher being the only one teaching CS in their school or district, and they might not be in a position to create or support a network of teachers. However knowing this may help providers better source additional supports (e.g., CSTA) to meet the needs of teachers that they themselves cannot meet.

7 CONCLUSION

Teachers are a crucial component of students learning CS, which has been shown to be empirically true in multiple research studies [42]. Effective CS PD is important to building teachers' level of CS CK and PCK so they are enabled to teach CS effectively and equitably, thus resulting in stronger academic achievement among their students. PD providers can measure this growth by leaning on the Novice to Expert Model, as well as understanding how teacher growth begins in undergraduate pre-service programs and continues throughout their professional career.

Our analysis examined the outcomes that U.S. CS PD providers evaluate, with an eye towards providing guidance to CS PD providers to determine its effectiveness. Since CS PD providers each have their own unique goals, a single evaluation instrument likely will not meet the needs universally. Our recommendations recognize certain important factors that CS PD providers may want to measure when determining the impact of their CS PD. Overall, the recommendations are a way to provide guidance in developing or choosing instruments for CS PD effectiveness.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. National Science Foundation under Grant Nos. 2039175.

REFERENCES

[1] Ahmet Oguz Akturk and Handan Saka Ozturk. 2019. Teachers' TPACK levels and students' self-efficacy as predictors of students' academic achievement. *International Journal of Research in Education and Science* 5 (2019). Issue 1.

[2] Alicia C Alonzo. 2007. Challenges of simultaneously defining and measuring knowledge for teaching. (2007).

[3] Daniel Alston, Jeff. Marshall, and Andrew Tyminski. 2017. Convincing Science Teachers for Inquiry-Based Instruction: Guskey's Staff Development Model Revisited. *Science Educator* 25 (2017). Issue 2.

[4] Judy Anderson and Deborah Tully. 2020. *Designing and Evaluating an Integrated STEM Professional Development Program for Secondary and Primary School Teachers in Australia.* https://doi.org/10.1007/978-3-030-52229-2_22

[5] Nalline Baliram and Arthur K. Ellis. 2019. The impact of metacognitive practice and teacher feedback on academic achievement in mathematics. *School Science and Mathematics* 119 (2019). Issue 2. https://doi.org/10.1111/ssm.12317

[6] A. Bandura. 1997. *Self-efficacy: The exercise of control.* W H Freeman/Times Books/ Henry Holt & Co.

[7] Eric R Banilower, P Sean Smith, Kristen A Malzahn, Courtney L Plumley, Evelyn M Gordon, and Meredith L Hayes. 2018. Report of the 2018 NSSME+. *Horizon Research, Inc.* (2018).

[8] Jürgen Baumert and Mareike Kunter. 2013. The effect of content knowledge and pedagogical content knowledge on instructional quality and student achievement. In *Cognitive activation in the mathematics classroom and professional competence of teachers.* Springer, 175–205.

[9] Patricia Benner. 1982. From novice to expert. *American Journal of nursing* 82, 3 (1982), 402–407.

[10] Edward M Bettencourt, Maxwell H Gillett, Meredith Damien Gall, and Ray E Hull. 1983. Effects of teacher enthusiasm training on student on-task behavior and achievement. *American educational research journal* 20, 3 (1983), 435–450.

[11] Moiz Bhai and Irina Horoi. 2019. Teacher characteristics and academic achievement. *Applied Economics* 51 (2019). Issue 44. https://doi.org/10.1080/00036846.2019.1597963

[12] David Blazar. 2016. *Teacher and teaching effects on students' academic performance, attitudes, and behaviors.* Ph.D. Dissertation.

[13] Rosemary Callingham, Colin Carmichael, and Jane M Watson. 2016. Explaining student achievement: The influence of teachers' pedagogical content knowledge in statistics. *International Journal of Science and Mathematics Education* 14, 7 (2016), 1339–1357.

[14] Computer Science Teachers Association. 2021. Quality PD Indicators. https://www.csteachers.org/Page/quality-pd-review-process

[15] National Research Council et al. 2001. *Classroom assessment and the national science education standards.* National Academies Press.

[16] Linda Darling-Hammond, Maria E Hyler, and Madelyn Gardner. 2017. Effective Teacher Professional Development. Research Brief. *Learning Policy Institute* (2017).

[17] Sloan Davis, Jason Ravitz, and Juliane Blazevski. 2018. Evaluating Computer Science Professional Development Models and Educator Outcomes to Ensure Equity. *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology, RESPECT 2018 - Conference Proceedings.* https://doi.org/10.1109/RESPECT.2018.8491716

[18] Yaron Doppelt, Christian D. Schunn, Eli M. Silk, Matthew M. Mehalik, Birdy Reynolds, and Erin Ward. 2009. Evaluating the impact of a facilitated learning community approach to professional development on teacher practice and student achievement. *Research in Science and Technological Education* 27 (2009). Issue 3. https://doi.org/10.1080/02635140903166026

[19] Stuart E Dreyfus. 2004. The five-stage model of adult skill acquisition. *Bulletin of science, technology & society* 24, 3 (2004), 177–181.

[20] David Dunning. 2011. The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in experimental social psychology.* Vol. 44. Elsevier, 247–296.

[21] Anna J Egalite and Brian Kisida. 2018. The effects of teacher match on students' academic perceptions and attitudes. *Educational Evaluation and Policy Analysis* 40, 1 (2018), 59–81.

[22] Roger D Goddard, Laura LoGerfo, and Wayne K Hoy. 2004. High school accountability: The role of perceived collective efficacy. *Educational Policy* 18, 3 (2004), 403–425.

[23] Thomas R Guskey. 2003. What makes professional development effective? *Phi delta kappan* 84, 10 (2003), 748–750.

[24] Karla Hamlen, Nigamanth Sridhar, Lisa Bievenue, Debbie K. Jackson, and Anil Lalwani. 2018. Effects of teacher training in a computer science principles curriculum on teacher and student skills, confidence, and beliefs. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education* 2018-January. https://doi.org/10.1145/3159450.3159496

[25] Ronald H Heck, Terry J Larsen, and George A Marcoulides. 1990. Instructional leadership and school achievement: Validation of a causal model. *Educational Administration Quarterly* 26, 2 (1990), 94–125.

[26] Wayne K Hoy and John W Hannum. 1997. Middle school climate: An empirical assessment of organizational health and student achievement. *Educational Administration Quarterly* 33, 3 (1997), 290–311.

[27] Helen H Hu, Cecily Heiner, Thomas Gagne, and Carl Lyman. 2017. Building a statewide computer science teacher pipeline. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education.* 291–296.

[28] Thomas J Kane, Daniel F McCaffrey, Trey Miller, and Douglas O Staiger. 2013. Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project. *Bill & Melinda Gates Foundation* (2013).

[29] Erdogan Kaya, Ezgi Yesilyurt, Anna Newley, and Hasan Deniz. 2019. Examining the impact of a computational thinking intervention on pre-service elementary science teachers' computational thinking teaching efficacy beliefs, interest and confidence. *Journal of Computers in Mathematics and Science Teaching* 38, 4 (2019), 385–392.

[30] Melanie M Keller, Knut Neumann, and Hans E Fischer. 2017. The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching* 54, 5 (2017), 586–614.

[31] Claudia Khourey-Bowers and Doris G Simonis. 2004. Longitudinal study of middle grades chemistry professional development: Enhancement of personal science teaching self-efficacy and outcome expectancy. *Journal of Science Teacher Education* 15, 3 (2004), 175–195.

[32] Lisa E Kim, Ilan Dar-Nimrod, and Carolyn MacCann. 2018. Teacher personality and teacher effectiveness in secondary school: Personality predicts teacher support and student self-efficacy but not academic achievement. *Journal of Educational Psychology* 110, 3 (2018), 309.

[33] Robert M Klassen and Ming Ming Chiu. 2010. Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of educational Psychology* 102, 3 (2010), 741.

[34] Kim Lange, Thilo Kleickmann, and Kornelia Möller. 2011. Elementary teachers' pedagogical content knowledge and student achievement in science education. In *ESERA-Conference, Lyon, France*. 5–9.

[35] Jihyun Lee and Valerie J Shute. 2010. Personal and social-contextual factors in K–12 academic performance: An integrative perspective on student learning. *Educational psychologist* 45, 3 (2010), 185–202.

[36] Monica M McGill, Leigh Ann DeLyser, Karen Brennan, Baker Franke, Errol Kaylor, Eric Mayhew, Kelly Mills, and Aman Yadav. 2020. Evaluation and assessment for improving CS teacher effectiveness. *ACM Inroads* 11, 4 (2020), 35–41.

[37] Muhsin Menekse. 2015. Computer science teacher professional development in the United States: a review of studies published between 2004 and 2014. *Computer Science Education* 25, 4 (2015), 325–350.

[38] Bismark Mensah and Eric Koomson. 2020. Linking Teacher-Student Relationship to Academic Achievement of Senior High School Students. *Social Education Research* (2020). https://doi.org/10.37256/ser.122020140

[39] Emmelien Merchie, Melissa Tuytens, Geert Devos, and Ruben Vanderlinde. 2018. Evaluating teachers' professional development initiatives: towards an extended evaluative framework. *Research Papers in Education* 33 (2018). Issue 2. https://doi.org/10.1080/02671522.2016.1271003

[40] Ahmad Mojavezi and Marzieh Poodineh Tamiz. 2012. The Impact of Teacher Self-efficacy on the Students' Motivation and Achievement. *Theory & Practice in Language Studies* 2, 3 (2012).

[41] National Center for Education Statistics. 2019. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d20/tables/dt20_214.10.asp

[42] Isaac M. Opper. 2019. *Teachers Matter: Understanding Teachers' Impact on Student Achievement*. RAND Corporation, Santa Monica, CA. https://doi.org/10.7249/RR4312

[43] Shivaun O'Brien, Gerry McNamara, Joe O'Hara, and Martin Brown. 2020. Learning by doing: evaluating the key features of a professional development intervention for teachers in data-use, as part of whole school self-evaluation process. *Professional Development in Education* (2020). https://doi.org/10.1080/19415257.2020.1720778

[44] Ursula Pieper and Jan Vahrenhold. 2020. Critical Incidents in K-12 Computer Science Classrooms-Towards Vignettes for Computer Science Teacher Training. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 978–984.

[45] Jason Ravitz, Chris Stephenson, Karen Parker, and Juliane Blazevski. 2017. Early lessons from evaluation of computer science teacher professional development in Google's CS4HS program. *ACM Transactions on Computing Education (TOCE)* 17, 4 (2017), 1–16.

[46] Susan J Rosenholtz. 1989. Workplace conditions that affect teacher quality and commitment: Implications for teacher induction programs. *The Elementary School Journal* 89, 4 (1989), 421–439.

[47] Ronny Scherer and Fazilat Siddiq. 2015. Revisiting teachers' computer self-efficacy: A differentiated view on gender differences. *Computers in Human Behavior* 53 (2015), 48–57.

[48] Alan H Schoenfeld. 2007. The complexities of assessing teacher knowledge. (2007).

[49] Khurram Shahzad and Sajida Naureen. 2017. Impact of Teacher Self-Efficacy on Secondary School Students' Academic Achievement. *Journal of Education and Educational Development* 4, 1 (2017), 48–72.

[50] Roger C Shouse. 1998. Restructuring's impact on student achievement: Contrasts by school urbanicity. *Educational Administration Quarterly* 34, 1_suppl (1998), 677–699.

[51] Lee Shulman. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard educational review* 57, 1 (1987), 1–23.

[52] Scott R Sweetland and Wayne K Hoy. 2000. School characteristics and educational outcomes: Toward an organizational model of student achievement in middle schools. *Educational Administration Quarterly* 36, 5 (2000), 703–729.

[53] Seçil Bal Taştan, Seyed Mehdi Mousavi Davoudi, Alfiya R. Masalimova, Alexandr S. Bersanov, Rashad A. Kurbanov, Anna V. Boiarchuk, and Andrey A. Pavlushin. 2018. The impacts of teacher's efficacy and motivation on student's academic achievement in science education among secondary and high school students. *Eurasia Journal of Mathematics, Science and Technology Education* 14 (2018). Issue 6. https://doi.org/10.29333/ejmste/89579

[54] Bruce W Tuckman and Thomas L Sexton. 1991. The effect of teacher encouragement on student self-efficacy and motivation for self-regulated performance. *Journal of Social Behavior and Personality* 6, 1 (1991), 137.

[55] Cynthia L Uline, Daniel M Miller, and Megan Tschannen-Moran. 1998. School effectiveness: The underlying dimensions. *Educational Administration Quarterly* 34, 4 (1998), 462–483.

[56] Mucella Ulug, Melis Seray Ozden, and Ahu Eryilmaz. 2011. The Effects of Teachers' Attitudes on Students' Personality and Performance. *Procedia - Social and Behavioral Sciences* 30 (2011), 738–742. https://doi.org/10.1016/j.sbspro.2011.10.144 2nd World Conference on Psychology, Counselling and Guidance - 2011.

[57] University of Southern California - Center for Urban Education. 2021. Equity Mindset. https://cue.usc.edu/about/equity/equity-mindedness/

[58] U.S. Census Bureau. 2019. Census Bureau Reports Nearly 77 Million Students Enrolled in U.S. Schools. https://www.census.gov/newsroom/press-releases/2019/school-enrollment.html

[59] Beverly Vaillancourt and Emmanuel Schanzer. 2018. In Search of Quality Professional Development. https://drive.google.com/file/d/1ztko8ePIVo7I-QQzPuUrCIB4-qItPc88/view

[60] Rebecca Vivian and Katrina Falkner. 2018. A survey of Australian teachers' self-efficacy and assessment approaches for the K-12 digital technologies curriculum. *Proceedings of the 13th Workshop in Primary and Secondary Computing Education* (2018).

[61] Herbert Ware and Anastasia Kitsantas. 2007. Teacher and collective efficacy beliefs as predictors of professional commitment. *The journal of educational research* 100, 5 (2007), 303–310.

[62] Aman Yadav and Marc Berges. 2019. Computer science pedagogical content knowledge: Characterizing teacher performance. *ACM Transactions on Computing Education (TOCE)* 19, 3 (2019), 1–24.

[63] Aman Yadav, Marc Berges, Phil Sands, and Jon Good. 2016. Measuring computer science pedagogical content knowledge: An exploratory analysis of teaching vignettes to measure teacher knowledge. In *Proceedings of the 11th Workshop in Primary and Secondary Computing Education*. 92–95.

[64] Aman Yadav, Alex Lishinski, and Phil Sands. 2021. Self-efficacy Profiles for Computer Science Teachers. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 302–308.

[65] Marjolein Zee and Helma MY Koomen. 2016. Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational research* 86, 4 (2016), 981–1015.

[66] Albert Zeggelaar, Marjan Vermeulen, and Wim Jochems. 2020. Evaluating effective professional development. *Professional Development in Education* (2020). https://doi.org/10.1080/19415257.2020.1744686

[67] Ninger Zhou, Ha Nguyen, Christian Fischer, Debra Richardson, and Mark Warschauer. 2020. High School Teachers' Self-efficacy in Teaching Computer Science. *ACM Transactions on Computing Education (TOCE)* 20, 3 (2020), 1–18.