Evaluation and Assessment for Improving CS Teacher Effectiveness

By Monica M. McGill, *Knox College & CSEdResearch.org*, Leigh Ann DeLyser, *CSforALL*, Karen Brennan, *Harvard University*, Baker Franke, *Code.org*, Errol Kaylor, *Knox College*, Eric Mayhew, *McGill University*, Kelly Mills, *Digital Promise*, and Aman Yadav, *Michigan State University*

As computing education becomes more integrated into K-12 classrooms, the need to concurrently develop teacher content knowledge and pedagogical content knowledge means that there are also time-sensitive needs for measuring the effectiveness of professional development (PD). However, what measures of assessment and evaluation exist? Where should funding be directed to fill those gaps? During a recent workshop, we set out to answer those questions. Our results indicate that there remain notable gaps in formative teacher assessment and evaluation, including within teacher impacts on equity, integrated computational thinking and computer science, human/curricular support structures, efficacy, and student impact factors.

Figure 1: As computing education becomes more integrated into K-12, there are also time-sensitive needs for measuring the effectiveness of professional development.

"DUKE ELLINGTON DNG 360" BY US DEPT OF EDUCATION IS LICENSED UNDER CC BY 2.0. TO VIEW A COPY OF THIS LICENSE, VISIT HTTPS://CREATIVECOMMONS.ORG/LICENSES/BY/2.0/

INTRODUCTION

To meet current demands for computing educators, K-12 teachers face an unprecedented task of learning computer science and computational thinking while simultaneously learning the pedagogical content knowledge needed to convey that knowledge successfully to students. Factors such as teacher self-efficacy, empowerment, and teacher support systems, have all been shown to directly impact academic achievement across a variety of other fields [18,27]. For example, teacher empowerment has been shown to strongly correlate to academic achievement in reading and mathematics [46] and collective teacher self-efficacy strongly correlates to reading, writing, and mathematics [22,47]. Teachers' attention to weak students, setting of performance targets, workload, administration of students' classroom assignments has also been shown to affect academic achievement [26].

As computing education becomes more integrated into K-12 classrooms [2,6,13,16,17], the need for concurrently developing teacher content knowledge and pedagogical content knowledge means that there are time-sensitive needs for measuring the effectiveness of professional development (PD). During the 2019 ACM International Conference on Education Research (ICER), a workshop was held with 26 evaluators, researchers, and other stakeholders to investigate evaluation and assessment needs across computing education [15]. Participants were grouped by their selected areas of interest: Primary, Secondary, Post-Secondary, and Teacher. This article's authors were in the teacher group, which had representation from across the United States and Canada and decades of collective experience in teacher PD, assessment and evaluation, and PD resource and capacity building. Our experiences include theoretical expertise with a deep understanding of practice, including former secondary computing education teachers, educational researchers, and computer scientists in academia. Our goal was to identify the current needs for evaluating teachers who teach computing education to enact an improvement science lifecycle, then consider how those needs could be met.

Our discussions borrowed methods from participatory action design and were guided by prompts like "What measures of assessment and evaluation do we currently use and have available to us?" but then dug deeper into questions like "What are our needs in our current research?" and "What views about assessment present barriers to its full integration into programs?" For each prompt, we placed our individual responses on sticky notes, followed by discussions of our responses (Figure 2). We then grouped the responses into categories to gain a broader sense and consider if the responses accurately reflected the state of the field. This led to a detailed overview of the gaps in these measures in K-12 computing education and steps the community can take to help fill these needs.

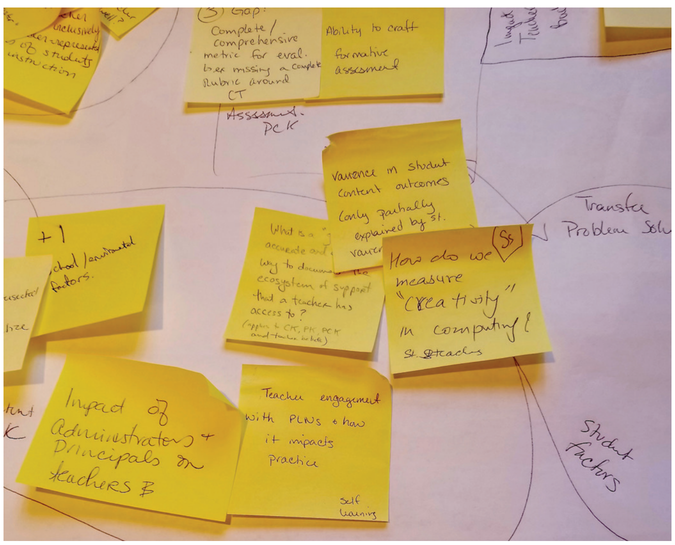

Figure 2: By answering and discussing prompts, we were able to tap into our individual expertise for formulating a larger picture of evaluation and assessment of teachers involved in teaching computer science and computational thinking.

This analysis is timely and relevant to those who want to understand teacher impact on computing education as it moves swiftly into primary and secondary schools, including funding bodies interested in engaging an improvement science model with the goal of improving student outcomes.

CURRENT ASSESSMENT AND EVALUATION MEASURES

Measures of student learning and engagement often take precedence over individual measures of teacher success and of measures of what happens in the classroom. Unfortunately, this creates a "closed box" effect where it is impossible to paint a full picture of not just what is happening with students, but how the environment contributes to those outcomes and can be improved to enable equitable outcomes for all learners. Researchers and evaluators can expand the view to include the intended curriculum (e.g., the curriculum as designed by in-district designers, packaged curriculum providers, and teachers) and the enacted curriculum (e.g., what happens in the classroom) [35,37]. These variations and their impact on learners need to be accounted for and discussed as an important part of assessments, since they can help explain the variability in student academic achievement across classrooms, schools, and educational settings. Teachers are at the heart of implementation, and therefore understanding teacher profiles, in addition to curricular manipulations, can provide additional information for researchers and implementers for program improvement.

In this article, we focus specifically on the characteristics of teachers that have been explored in CS education research, as well as discussed in the workshop described above. CS education implementation efforts have focused on teacher PD, and therefore a significant portion of publications includes examples of both qualitative [23] and quantitative measures [33] of teachers' content knowledge and beliefs. Additionally, national studies give us curricular independent examples of teacher assessment [3,20]. We defined the four focus areas that follow.

Content knowledge includes micro-credentials, teacher reflection, surveys, Scantron tests such as "Which of the following lines of code complete the task," and assessment of projects.

Pedagogical content knowledge includes assessing teachers via surveys (e.g., PCK surveys, Vignettes, use of reform oriented practices), observation/interviews (e.g., assessing how students decompose algorithms, use parameters/arguments, classroom observations, interviews to understand teacher practices, discuss creative design aspirations, and classroom observations-intentions), and artifacts (design journals, collecting examples of practice, revised aspirations, lesson plans).

Efficacy/beliefs includes self-reported evaluation measures including attitudes/self-efficacy, teacher self-confidence on individual units, use of instructional strategies in the curriculum, teaching practices, community support, engagement, and equity practices. Group members have also evaluated teacher beliefs of students' growth mindset, belief in student outcomes, and dispositions toward equity using identity awareness/blindness and dichotomous statements.

Program evaluation includes assessing administrators or planning of support for initial PD alignment, ongoing learning, concerns based model of adoptions (self, task, impact), teacher leadership advocacy, participation in research, which students are engaged in learning, and the identification of challenges that teachers face and their desired level of support.

These four categories are evident in other teacher quality evaluations. The National Survey of Science and Mathematics Education (NSSME) similarly divide their measures into Teacher Background and Beliefs, Professional Development, Courses Delivered, Instructional Decision Making, Objectives, and Activities, Instructional Resources, and Factors Affecting Instruction [3]. The CS Education community can draw more background and experience from broader education and teacher effectiveness literature in thinking about what factors may indicate teacher quality and ultimately impact student outcomes.

GAPS IN ASSESSMENT AND EVALUATION MEASURES

The development of CS research assessments with evidence of validity is still an open research area [48]. We identified several gaps in measures for assessing and evaluating teacher factors impacting student academic achievement, categorizing them as Teacher Impacts on Equity (with a subcategory of Culturally Relevant/Responsive Instructional Practice), Integrated CS, Human/Curricular Support Structures, PCK-Efficacy, and Student Impact Factors.

Given their importance, each of these is described in the following subsections. For each category, we discussed and came to a consensus on its Possibility of Design (Easy, Medium, or Difficult), Possibility of Implementation (Easy, Medium, or Difficult), and Priority Level (Low, Medium, or High) (see Figure 3).

TEACHER IMPACTS ON EQUITY

One of the goals for broadening participation is to provide every student the opportunity to learn CS [14]. Teachers play a crucial role in creating and sustaining the environment needed to make this happen [29]. Acknowledging that teachers often teach students who do not look like them, their personal experiences are not relevant in either a modern pedagogical environment (cannot reflect on their own school) or in the cultural context where they teach. With broadening participation goals, it is important that we acknowledge that this is not only critically important, but that even experienced educators may lack the competence in equity discussions as many teachers come from a STEM background and are not used to deep discussions of culture and ethics. Therefore, there is a need for measurements to assess teacher knowledge of how to engage all students equitably in their classrooms. This includes the ability to measure how well the community facilitates PD experiences designed to prepare teachers to bring CS to students in culturally relevant ways and how well PD prepares teachers to leverage community-based resources.

Teachers make instructional and structural decisions every day that impact equitable outcomes, including lesson presentation, skill practice, questioning techniques, discipline, and how the worth and dignity of each student is attained [25,42]. Structural decisions may include recruitment strategies and targeted populations, knowledge of resources to engage under-represented youth in considering CS classes, and appropriate images (classroom decorations) [4], and promising practices for messaging about the content within the course. Instructional decisions happen both in the planning and delivery of course content [25,41,42], and the ability for teachers to consider equity in those moments is also an important skill that needs to be assessed to evaluate teacher efficacy.

Promising practices that reinforce equitable outcomes are grounded within high quality instruction [29]. The NSSME labels impactful practices as reform-oriented practices and include collaborative environments that make use of real-world problems as an example of a reform-oriented practice that is aligned with promising practices for encouraging equity [3].

As an important subcategory of teacher impacts on equity, measurements are needed for assessing and evaluating the ways in which teachers are prepared to adopt lessons to best meet cultural and community values, challenges, and interests of their students. This practice is not about just borrowing "diversity" lessons, but rather developing teachers' capacity for richly crafting new lessons or adapting lessons to meet their own individual student populations. This corresponds to research in other fields where pre-service teachers are trained and evaluated [44], the impact of PD on teachers for culturally relevant practices is explored [39], and the examination of teachers practice of including culturally relevant practices in their everyday teachings [7,8]. Specifically, we declare a need for assessment and evaluation methods to measure several factors in computing education, including the qualifications of teachers, engagement of underrepresented populations, teacher expertise (including development and evolution over time), and others shown in Figure 4.

We rated this category's design possibility as easy due to the ability to observe practice in a classroom based upon promising practices in the CS education literature. Implementation is *difficult*, since it takes significant resources to do this at scale. Even so, creating measures of teacher impacts on equity is *high priority*.

	POSSIBILITY OF DESIGN	POSSIBILITY OF IMPLEMENTATION	PRIORITY LEVEL
Teacher Impacts on Equity	Easy	Difficult	High
Integrated CT/CS	Very Difficult	Medium	High
Human/Curricular Support Structures	Medium	Easy	Medium to High
PCK-Efficacy	Difficult	Difficult	Low
Student Factors	Easy	Medium	Very High

Figure 3: By identifying the evaluation and assessment measures most needed and achievable, funding and resources can be directed to these needs.

Gaps in Assessment and Evaluation: Equity and Student Outcomes

- Level and Impact of Highly Qualified (Master) Teachers
- Inclusive engagement of underrepresented student populations
- Classroom observation at scale, classroom dynamics, equitable outcomes
- Differences in outcomes for under-represented minorities and gender
- Factors that influence teachers adoption of practice
- Development/evolution of teacher expertise over time
- Culturally relevant (CR)/responsive instructional practices

Figure 4: Equity and student outcomes are important factors in knowing whether teachers are prepared to impact all students.

INTEGRATED CT/CS

There are currently many initiatives focused on integrating Computational Thinking (CT) and Computer Science (CS) into other knowledge areas, such as mathematics, biology, and other STEM fields [43,49]. Our field has developed initial understandings about how some CS practices may impact student learning in specific disciplines (e.g., using a computational model to explore a physical phenomenon). Within this category, we seek to understand how a broader subset of CS practices (e.g., abstraction, debugging, pattern recognition) may impact content learning, particularly applied to disciplines in which computer science is not traditionally integrated. We also recommend the further exploration of how content teachers can facilitate CS practices in ways that build students' knowledge and efficacy in computing.

This category requires measurements in both PCK and Content Knowledge (CK), including measuring the impact that CT/CS integration has on other academic achievement and teaching in other disciplines and the teacher-facilitation strategies that strengthen CT/CS integration. While some CS curricula have been designed to integrate concepts from CS/CT into other academic areas, there is little research into the impact that teacher preparation and practice can have on integrated approaches.

We rated this category as high priority. However, it is exceedingly difficult to design such measures due to their complexity and layers of content and pedagogy being enacted. Implementation is challenging as well due to the difficulties in aligning content within other subdomains and reaching agreement among experts for an assessment or evaluation measure's usage.

HUMAN/CURRICULAR SUPPORT STRUCTURES

There are a broad range of rich curricula and PD programs, communities of practice, and professional learning networks for CT/CS teachers, especially compared to a few years ago. There is also a variable amount of development of programs that affect systems and environments in which teachers are situated (e.g., policy, principal support, school factors). Evaluating these structures and ecosystems is important to understand how they impact teacher development and growth (e.g., PCK), mastery, and leadership, as well as determining which are most critical with respect to academic achievement in computing education (see Figure 5).

Gaps in Assessment and Evaluation: Human/Curricular Support Structures

- Factors most influential in changing PCK
- Knowledge of standards and mapping of curriculum to standards
- Knowledge of computational thinking
- Effect of paired teacher
- Impact of colleagues on teachers
- Impact of teacher evaluation systems with CS education practice
- Impact of teacher agency
- Impact of school/environment factors on teachers
- Impact of administration on teachers
- Impact of community on teachers
- Impact of teacher demographic profiles on student learning and self-efficacy
- Impact of teacher-accessible resources reflecting the ecosystem of support
- Impact of teacher engagement with Professional Learning Networks (PLN) on practice
- Intersection between teacher evaluation and systems to support effective teaching

Figure 5: Human and curricular support structures have a measurable impact on academic achievement on K-12 learners.

Education implementation research provides a theoretical framing for the hypothesis that the structures that surround teachers, both human and curricular, impact teacher's performance and development. Some key factors that have only been briefly explored in CS education include instructional coherence [10], distributed leadership [45], teacher social networks [11], and more. Each of these structural supports (or barriers) can contribute to the overall efficacy of teachers and should be considered in any evaluation of teacher development.

Measures of human and curricular support structures and their impact on academic achievement that are still needed include those that measure teacher agency [8,36,38], teacher engagement in communities of practice [1,12], and knowledge of standards with the ability to map standards to curriculum [34,40]. Figure 3 defines a list that will enable the research and evaluation community to move forward in developing these measures.

We rated this category has a medium to high priority. However, the possibility of its design is medium due to the category requiring the measurement of the substantive impact of the human/curricular support structures. The possibility of implementing these measures is easy.

PCK EFFICACY

The Pedagogical Content Knowledge (PCK) Efficacy category emerged as the combination of teacher PCK and the implications of that for teacher practice. As shown in other subject areas, those implications can have direct impacts on practice (being savvy to Gaps in Assessment and Evaluation (Human/Curricular Support Structures) teach) and academic achievement [5,21,24] as well as indirect impacts in terms of transferring CS to integrated scenarios. In this category PCK is not the directly assessed outcome, but a contributor to another measurement. Several measures of assessment and evaluation for PCK-Efficacy are still needed in K-12 computing education, including the following.

- Impact of the teacher self-efficacy on PCK for in-service subject changers
- Impact of tech "savviness" on teaching (Technological Pedagogical Content Knowledge (TPACK), Computer Science Pedagogical Content Knowledge)
- Knowledge needed integrate CS into content areas
- Impact of PCK on teaching a lesson on computing

We rated the possibility of design measures in PCK Efficacy as difficult, the possibility of implementation as difficult, and the priority level as low.

STUDENT IMPACT FACTORS

The assessment of student factors was derived from the need to measure student factors that are reflective of teacher implementation of curriculum [28]. Assessment of student factors also can be used to measure how teachers' implementation of these factors influence their own practice [19,30,31]. These measures allow teachers to reflect on their own practices and how they drive or influence the impact of their teaching and opportunities for learning. Gaps for assessing and evaluating students to inform and improve teachers' PCK include the following.

- Transfer of knowledge
- Synthesis of techniques (e.g., blocks into solving other tools)
- Creativity
- Ability to craft assessment on creativity/transfer
- Ability to craft formative assessment
- Ability to determine appropriateness of assessment
- Ability to craft summative assessment
- Variance in student content outcomes (though only partially explained by student variance)

We rated this category as easy to design, medium for implementation, and of extremely high priority.

RECOMMENDATIONS

With respect to assessment, then, what are the changes that are most needed? And what recommendations can we provide to the larger community to start to enable the creation of these measures?

GUIDING VALUES FOR ASSESSMENT

Rather than viewing assessment and evaluation as a single step that happens at the end of an intervention, we encourage readers to view it more holistically. We invite the community to consider the following guiding values when assessing and evaluating (also see Figure 6).

We communicate our community's values through assessment. What matters most to the community should be driving those program elements that should be assessed. The goals, as set forth by the community, must be specifically mapped to assessment. Community involvement includes the input from teachers through the shared language and understanding of what is important as well as the "lived experiences" of students. These then directly translate into what needs to be assessed.

We use assessment to peel back layers of those factors that impact student learning and to discover what is happening in the classrooms. Assessment can help pave the way for teacher growth and student learning in the classroom.

We see further need to shift from the more ad hoc practice of assessment to one that emphasizes and upholds improvement science. We cannot improve what we do not measure and what we are unable to understand. Developing assessments based on theory and then assessing the results according to theory will help move the practice from a more ad hoc practice to one that has a solid foundation.

To achieve CS for All, we must work to change teachers' belief systems about equity in addition to how these beliefs can be integrated into the classroom. Though many of these practices are in place in many classrooms, moving these practices into all classrooms in all schools will have a tremendous impact on achieving CS for All.

To achieve full adoption of these principles, many cultural and attitude shifts must occur within the field. For example, K-12 teachers and administrators have some input and control over PD design, standards, and policies, which is a start in understanding community values and as well as the enacted classroom curriculum. However, they have less control over funding, the design of coherent pathways, and teacher preparation programs.

These values and their implementation require further consideration. For example, how do we ensure that the considerations of others within the community, including the perspectives of learners (students), are included in setting these values? What, then, do we need to do to ensure that their agency is fully optimized in this process given their constraints? As the field

Guiding Values for Assessment	We communicate our community's values through assessment. What matters most to the community should be driving those program elements that should be assessed.	We use assessment to peel back layers of those factors that impact student learning and to discover what is happening in the classrooms.	We see further need to shift from the more ad hoc practice of assessment to one that emphasizes and upholds improvement science. We cannot improve what we do not measure nor understand.	To achieve CS for All, we must work to change teachers' belief systems about equity in addition to how these beliefs can be integrated into the classroom.	To achieve full adoption of these principles, many cultural and attitude shifts must occur within the field.

Figure 6: Applying growth mindset principles to the assessment of teachers is critical in ensuring that the community's values are incorporated, and teachers are given the support needed to grow.

continues to evolve and grow, these are important questions in addressing academic achievement in a comprehensive manner that reflects the values of everyone within the community.

KEY ACTIVITIES TO UNDERTAKE

There are several key activities that the community (including sponsors and funding organizations) can undertake to fill these needs. Within each, there are multiple opportunities for stakeholders worldwide to enable researchers and evaluators to address these many needs.

Involve and Catalyze Teachers. Involve and catalyze teachers to get their input (with guidance and discretion). This can be achieved through engaging teachers and other stakeholders in the design process for indicators of high quality teaching/PCK, including out of school hours instructors as part of informing assessment, areas for growth, assessing teachers not only around PD, but how intended curriculum moves into practice (enacted curriculum), and studying the many factors that affect student learning.

Provide Feedback to Funding Bodies. Craft a message that states the research needs in the community followed by discussions with funding bodies (e.g., program officers at government institutions and agencies, private funding sponsoring organizations) for meeting the assessment/evaluation needs of for the research within the K-12 community. This includes efforts to bridge the gap between results and assessment, build and disseminate assessment, and ensure that grants focus specifically on assessment, rather than as a supplemental option for other grants

Inform the CS Ed Community. Bring our knowledge of assessment and evaluation within the greater CS Ed Community by a variety of means.

- Hold continuous birds of a feather around teacher assessment (SIGCSE, other conferences/venues) to collect exemplars and factors
- Hold sessions at teacher conferences (e.g., ISTE, CSTA) to keep moving these items forward
- Create a computer science education research conference with funding by sponsoring agencies
- Create and maintain resources related to assessment and evaluation
- Continue the work of CSEdResearch.org as it seeks to grow a repository of assessments for teachers, students, and other factors influencing learning [32]
- Build instruments and frameworks to help stakeholders design PD for classification/categorizing the capacity of what they are building
- Evaluate the evaluation instruments and assessment methods, including determining their usefulness

Improve professional development. Improve PD for K-12 teachers in CT/CS.

- Create a better synthesis of the quality of the mastery of CS teachers' education, such as the branching tree of cognitive theories for CS education from Carnegie-Mellon University's Simon Initiative [9]
- Build PD based upon theory and assessment of theory
- Identify how to implement CS curricula by identifying and making individual exemplars for the wider community

CONCLUSION

We consider this to be the start of the conversation on identifying the needs for formative assessment of K-12 teachers in computing education and answering many important questions embedded in the process. We further recommend that priorities are based on the possibility of design, possibility of implementation, and priority level. Human/Curricular Support Structures, Teacher Impacts on Equity, and Student Factors rated high in the prioritization matrix. Integrated CT/CS is also important, but its design possibility was rated as exceedingly difficult due to its complexity. PCK Efficacy is both difficult to design and implement and, with a low priority level, we did not rate this factor as important as others.

Carrying this forward into the larger ACM community engaged in broadening participation and CS education efforts, we invite the computing education research and evaluation community to consider how to prioritize resources and enable the development of these instruments and assessment models—but within the context of improvement science. As computing education continues to ramp up in the K-12 space, the process of gathering metrics to enable teachers to reach all students will not be easy, particularly at scale. Focusing on small and incremental steps will enable this process to unfold and build a solid foundation for computing education for years to come. ❖

Acknowledgements

This material is based upon work supported by the U.S. National Science Foundation under Grant Nos. 1625005, 1625335, 1757402, 1745199, and 1933671. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Au, K. H. Communities of practice: Engagement, imagination, and alignment in research on teacher education. *Journal of Teacher Education* 53, 3 (2002), 222–227.
2. Australian Curriculum, Assessment and Reporting Authority (ACARA). Australian Curriculum: Digital Technologies; http://www.australiancurriculum.edu.au/. Accessed 2020 October 24.
3. Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., and Hayes, M. L. *Report of the 2018 NSSME+*; http://horizon-research.com/NSSME/2018-nssme/research-products/reports/technical-report. Accessed 2020 Oct 26.
4. Barker, L. J., and Garvin-Doxas, K. Making visible the behaviors that influence learning environment: A qualitative exploration of computer science classrooms. *Computer Science Education* 14, 2 (2004), 119–145.
5. Baxter, J. A., and Lederman, N. G. Assessment and measurement of pedagogical content knowledge. In *Examining Pedagogical Content Knowledge*. Springer, 1999, 147–161.
6. Bell, T., Andreae, P., and Robins, A. A case study of the introduction of computer science in NZ schools. *ACM Transactions on Computing Education* 14 (2014), 1–31.
7. Benson, B. E. Framing culture within classroom practice: Culturally relevant teaching. *Action in Teacher Education* 25, 2 (2003), 16–22.
8. Campbell, E. *Teacher Agency in Curriculum Contexts*, 2012.
9. Carnegie Mellon University's Simon Initiative; https://Learnlab.org. Accessed 2020 October 24.
10. Cobb, P., Jackson, K., Henrick, E. C., Smith, T. M., team, M., et al. *Systems for Instructional Improvement: Creating Coherence from the Classroom to the District Office*. Harvard Education Press Cambridge, MA, 2018.
11. Coburn, C. E., and Russell, J. L. District policy and teachers' social networks. *Educational Evaluation and Policy Analysis* 30, 3 (2008), 203–235.
12. Coburn, C. E., and Stein, M. K. Communities of practice theory and the role of teacher professional community in policy implementation. *New Directions in Education Policy Implementation: Confronting Complexity* (2006), 25–46.

13. Code.org. State of computer science education; https://advocacy.code.org/2019_state_of_cs.pdf. Accessed 2020 October 24.
14. CSforAll; https://www.csforall.org/ Accessed 2020 October 24.
15. Decker, A. and McGill, M.M. 2019. Student Learning: Creating, Refining, and Promoting Evaluation and Research Across Computing Education Workshop. ACM ICER 2019.
16. Department for Education. The National Curriculum in England. Department for Education Government of UK, Crown, Cheshire, 2013.
17. Directorate for Learning and Assessment Programmes. SECsyllabus (2019): Computing. https://www.um.edu.mt/data/assets/pdf_file/0017/292310/SEC09.pdf. Accessed 2020 October 24.
18. Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., and Beechum, N. O. Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance–A Critical Literature Review. ERIC, 2012.
19. Fuchs, L. S., and Fuchs, D. Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children* 53, 3 (1986), 199–208.
20. Gal-Ezer, J., and Stephenson, C. Computer science teacher preparation is critical. *ACM Inroads* 1, 1 (2010), 61–66.
21. Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D., and Stuhlsatz, M. A. Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education* 41, 7 (2019), 944–963.
22. Goddard, R. D., Hoy, W. K., and Hoy, A. W. Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal* 37, 2 (2000), 479–507.
23. Goode, J., Margolis, J., and Chapman, G. Curriculum is not enough: the educational theory and research foundation of the exploring computer science professional development model. In *Proceedings of the 45th ACM technical symposium on Computer science education* (2014), ACM, 493–498.
24. Keller, M. M., Neumann, K., and Fischer, H. E. 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.
25. Kemp, L., and Hall, A. H. Impact of effective teaching research on student achievement and teacher performance: Equity and access implications for quality education (1992).
26. Kimani, G. N., Kara, A. M., and Njagi, L. W. Teacher factors influencing students' academic achievement in secondary schools in Nyandarua county, Kenya. *International Journal of Education and Research* 1, 3 (2013), 1–14.
27. Lee, J., and Shute, V. J. Personal and social-contextual factors in k–12 academic performance: An integrative perspective on student learning. *Educational Psychologist* 45, 3 (2010), 185–202.
28. Lee, J., and Stankov, L. Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences* 65 (2018), 50–64.
29. Lewis, C. M., Shah, N., and Falkner, K. Equity and diversity. In the *Cambridge Handbook of Computing Education Research*, S. Fincher and A. Robins, Eds. Cambridge University Press, Cambridge, UK, 2019, 481–510.
30. Little, O., Goe, L., and Bell, C.A practical guide to evaluating teacher effectiveness. *National Comprehensive Center for Teacher Quality* (2009).
31. Mathers, C., and Oliva, M. Improving instruction through effective teacher evaluation: Options for states and districts. TQ research & policy brief. *National Comprehensive Center for Teacher Quality* (2008).
32. McGill, M. M., and Decker, A. Computer science education resource center (2019).
33. Milliken, A., Cody, C., Catete, V., and Barnes, T. Effective computer science teacher professional development: Beauty and joy of computing 2018. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education* (2019), 271–277.
34. Niess, M. L., Ronau, R. N., Shafer, K. G., Driskell, S. O., Harper, S. R., Johnston, C., Browning, C., Özgün-Koca, S. A., and Kersaint, G. Mathematics teacher TPACK standards and development model. *Contemporary Issues in Technology and Teacher Education* 9, 1 (2009), 4–24.
35. Nolet, V., and McLaughlin, M. J. *Accessing the General Eurriculum: Including Students with Disabilities in Standards-based Reform*. Thousand Oaks, CA: Corwin Press, Inc., 2000.
36. Palmer, D., and Martínez, R. A. Teacher agency in bilingual spaces: A fresh look at preparing teachers to educate latina/o bilingual children. *Review of research in Education* 37, 1 (2013), 269–297.
37. Porter, A. C., and Smithson, J. L. Defining, developing, and using curriculum indicators. CPRE research reports, 2001.
38. Priestley, M., Edwards, R., Priestley, A., and Miller, K. Teacher agency in curriculum making: Agents of change and spaces for manoeuvre. *Curriculum Inquiry* 42, 2 (2012), 191–214.
39. Roehrig, G. H., Dubosarsky, M., Mason, A., Carlson, S., and Murphy, B. We look more, listen more, notice more: Impact of sustained professional development on head start teachers' inquiry-based and culturally relevant science teaching practices. *Journal of Science Education and Technology* 20, 5 (2011), 566–578.
40. Roehrig, G. H., Moore, T. J., Wang, H.-H., and Park, M. S. Is adding the E enough? Investigating the impact of k-12 engineering standards on the implementation of stem integration. *School Science and Mathematics* 112, 1 (2012), 31–44.
41. Sadker, M. and Sadker, D. Sexism in the classroom: From grade school to graduate school. *The Phi Delta Kappan* 67, 7 (1986), 512–515.
42. Sammons, P., and Bakkum, L. Effective schools, equity and teacher efficacy: A review of the literature. Profesorado, Revista de Currículum y Formación del Profesorado 15, 3 (2011), 9–26.
43. Settle, A., Goldberg, D. S., and Barr, V. Beyond computer science: computational thinking across disciplines. In *Proceedings of the 18th ACM conference on Innovation and technology in computer science education* (ACM, 2013), 311–312.
44. Sleeter, C. E., and Owuor, J. Research on the impact of teacher preparation to teach diverse students: The research we have and the research we need. *Action in Teacher Education* 33, 5-6 (2011), 524–536.
45. Spillane, J. P. *Distributed leadership*, vol. 4. (John Wiley & Sons, 2012).
46. Sweetland, S. R., and Hoy, W. K. School characteristics and educational out-comes: Toward an organizational model of student achievement in middle schools. *Educational Administration Quarterly* 36, 5 (2000), 703–729.
47. Tschannen-Moran, M., and Barr, M. Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools* 3, 3 (2004), 189–209.
48. Vahrenhold, J., Cutts, Q., and Falkner, K. Schools (k-12). In the *Cambridge Handbook of Computing Education Research*, S. Fincher and A. Robins, Eds. Cambridge University Press, Cambridge, UK, 2019, 547–583.
49. Weintrop, D., Beheshti, E., Horn, M. S., Orton, K., Trouille, L., Jona, K., and Wilensky, U. Interactive assessment tools for computational thinking in high school stem classrooms. In *International Conference on Intelligent Technologies for Interactive Entertainment* (2014), Springer, 22–25.

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

Leigh Ann DeLyser
CSforALL
New York, NY
leighann@csforall.org

Karen Brennan
Harvard University
Cambridge, MA
karen_brennan@gse.harvard.edu

Baker Franke
Code.org
Chicago, IL
baker@code.org

Errol Kaylor
Knox College
Galesburg, IL
ebkaylor@knox.edu

Eric Mayhew
McGill University
Montreal, Quebec
eric.mayhew@mail.mcgill.ca

Kelly Mills
Digital Promise
San Mateo, CA
kmills@digitalpromise.org

Aman Yadav
Michigan State University
East Lansing, MI
ayadav@msu.edu

DOI: 10.1145/3410478

©2020 ACM 2153-2184/20/12 $15.00