Exploring K-8 Teachers' Preferences in a Teaching Augmentation System for Block-Based Programming Environments

Minji Kong
mkong@udel.edu
University of Delaware
Newark, Delaware, USA

Matthew Louis Mauriello
mlm@udel.edu
University of Delaware
Newark, Delaware, USA

Lori Pollock
pollock@udel.edu
University of Delaware
Newark, Delaware, USA

ABSTRACT

Multiple disciplines have taken interest in investigating and using teaching augmentation (TA) tools that are designed to support teachers' pedagogical capabilities during classroom activities. TA systems can take various forms (e.g., dashboards, ambient displays). However, research on TA systems that complement K-8 teachers' in-class when their students are learning to program in block-based programming environments (BBPEs) is nascent. For a TA system to positively impact teaching practices, the system's design should be informed by a strong understanding of its stakeholders' preferences. Through 10 semi-structured interviews with and 37 anonymous survey responses from K-8 teachers, we identify respondents' preferences for potential BBPE TA systems. To put their preferences into context, we also describe how respondents typically teach programming using a BBPE and monitor students' progress. Our mixed-methods approach reveals how TA systems could best target teachers' attention level when teaching using BBPEs and assist in interpreting students' behaviors while learning to code. Using these findings, we identify directions for future TA systems to best assist teachers in making data-driven instructional decisions and meeting students' learning needs.

CCS CONCEPTS

• Social and professional topics → K-12 education; **• Human-centered computing → Empirical studies in HCI**.

KEYWORDS

teaching augmentation tools, block-based programming environments, teaching pedagogy

ACM Reference Format:

Minji Kong, Matthew Louis Mauriello, and Lori Pollock. 2022. Exploring K-8 Teachers' Preferences in a Teaching Augmentation System for Block-Based Programming Environments. In *Koli Calling '22: 22nd Koli Calling International Conference on Computing Education Research (Koli 2022), November 17–20, 2022, Koli, Finland*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3564721.3564725

Koli 2022, November 17–20, 2022, Koli, Finland

ACM ISBN 978-1-4503-9616-5/22/11…$15.00
https://doi.org/10.1145/3564721.3564725

1 INTRODUCTION

Given that easily accessible computing education (CEd) has become a globally recognized need, researchers have designed and introduced block-based programming environments (BBPEs) with the aim of making programming and the computing field more approachable for all. BBPEs, including Scratch [64], iSnap [60], and Blockly [31], offer drag-and-drop graphical block representations of programming instructions for users to create a working program without syntax errors. These blocks allow users to focus on computing concepts and programming with reduced cognitive load [19]. Given the benefits of BBPEs for novice programmers, they have gradually grown in popularity since their introduction. Notably, Scratch has grown a reputation as one of the most popular BBPEs, with 37 million registered users ages 8-16 and over 27,000 supporting educators [64].

BBPEs alone, however, cannot provide high-quality computing instruction to all students and address challenges teachers may face. Research shows that computing pedagogy plays an important role in effectively introducing computing at the K-12 level [80]. In particular, recent literature indicates that the pedagogical needs that are most common among K-12 computing teachers include obtaining guidance on how to best implement student-centered teaching practices [67]. This need aligns with teaching challenges that accompany the increased demand for personalized learning [56], such as keeping track of students' learning processes and adapting teaching practices accordingly.

One way that K-12 computing teachers can obtain guidance on student-centered practices is through technologies that are designed to augment and complement teachers' pedagogical practices during ongoing in-class activities. Examples of such technologies, or *teaching augmentation* (TA) systems [13], include learning analytics dashboards, ambient displays, and wearables (e.g., [12, 25, 35, 38]). Although there has been rapid growth and interest in the range of TA system designs, this has not been the case in the context of BBPEs like Scratch. To the best of our knowledge, there are no Scratch-focused TA systems designed to support K-12 computing teachers. Furthermore, work on real-time analytics for personalized classrooms has been predominantly in university contexts [66].

In this paper, we investigate K-12 computing teachers' needs in a potential TA system specific to Scratch. In particular, we focus on the experiences and needs of K-8 computing teachers as Scratch is primarily designed for users age 8-16, and research has shown that students' interest in exploring Computer Science is often determined in middle school [64]. As early work, our research questions are exploratory and include: *What current unmet teaching needs do K-8 teachers using Scratch identify that augmentation systems could assist with?* And, *What augmentation forms do they prefer most?*

To address these questions, we conducted a need-finding study that involved a mixed-methods approach. Specifically, we first conducted semi-structured interviews with ten K-8 computing teachers, all with prior experience teaching programming using Scratch. We then complemented this study with data from a survey study of teachers ($n = 37$ respondents) designed based on the interview responses to validate our findings with a broader audience. This approach fit best in the context of our research as little is known about K-8 teachers' thoughts and preferences for augmentation, and our qualitative interview data are used to explore the topic and construct a survey instrument to gather relevant data from an additional sample [51].

Our findings highlight challenges and opportunities for researchers to consider when designing a Scratch-based (or similar BBPE) TA systems to ensure that the contributions benefit K-8 teachers and ultimately, the students, in computing classrooms. As a result, the main contributions of this work include: (i) insights about teachers' preferences for TA systems involving Scratch and (ii) design recommendations for TA systems which we believe will be useful for researchers and practitioners.

2 BACKGROUND AND RELATED WORK

Our research on exploring the concept of future Scratch and other BBPE TA system designs is motivated by the potential of such systems to help create a more positive affective orientation and interest among students toward the computing field at early stages of their education careers. Self-efficacy—or an individual's beliefs in their capabilities to complete a task [17], which, in our context, is BBPE programming—is one of the fundamental factors that affect one's attitude towards a domain; this has been illustrated in closely related areas such as scientific education [26]. Prior research in CEd has shown that students' self-efficacy predicted their outcomes and engagement in a Computer Science (CS) course [46], and relates significantly to their CS career orientation [9].

Unfortunately, there exists variance in levels of self-efficacy among students engaging with computing materials. Students of groups historically underrepresented in the computing field (e.g., gender, race, ethnicity) are more likely to face negative experiences, such as not feeling welcomed in or disconnected from computing [43], which impacts their self-efficacy toward the domain [75]. This challenges the need for the computing field to include and be respectful of all communities in order for computing innovations to align with the needs of society's demographics. With future BBPE TA system designs, we aim to assist teachers in better observing and meeting diverse students' varying learning needs, helping as many students as possible feel confident when learning about the computing field via BBPEs and ultimately have increased sense of self-efficacy toward computing.

Researchers have introduced TA systems across several disciplines, such as Human-Computer Interaction (HCI) [11, 25], learning analytics [50], and the learning sciences [74]. System designs come in different forms (an illustration of example designs is shown in Figure 1). For instance, *learning analytics dashboards* (e.g., [49, 50, 71]) provide teachers real-time information on how students are progressing and performing in their classes with the aim of enhancing teachers' in-class awareness. Alternative designs, inspired by the HCI notion of *calm technology* [79], make augmentation available at the periphery of teachers' attention rather than dashboards needing teachers' focal attention. This includes *ambient displays*, which can extend both teachers' and students' in-class situational awareness. For example, Lernanto indicates each student's learning pace through arrays of LEDs mounted on a wall, giving teachers background awareness that can help center their teaching practices around students' exhibited learning needs [72]. Similarly, CawClock uses visuals and soundscapes to enhance awareness of in-class activities [16].

There also exist *distributed* TA systems, which introduce *ubiquitous computing* in classrooms [78]. Distributed digital lamps, for instance, are physically placed in classrooms and are designed to provide ambient information about a class' activities (e.g., [11]). Lantern places ambient lamps at student desks to create a classroom orchestration system, with which work progress and help requests of university student teams can be depicted through signals such as pulse rate and color [10]. Firefly systems are similar in that they create non-verbal communication channels between teachers and students via the color shown by each student's or team's lamp [11, 25].

Several projects have also explored the idea of augmenting teachers' abilities with *wearables* (e.g., [63]), such as smartwatches and earpieces, some of which are designed to synchronously coach novice teachers through a remote observer. In addition, researchers have looked into the possibilities of using smart glasses (e.g., [38]). With wearables, teachers would be augmented without their attention diverted away from their classrooms.

BBPE-based research related to augmentation is evolving but still in its early stages. Many efforts focus on using learning analytics to study patterns of students' programming behaviors (e.g., construction, compilation, debugging) [20, 21, 23]. For example, Piech et al. [58] collected data from thousands of Code.org users as they worked on *Hour of Code* block-based programming activities, with the overall goal of predicting the types of hints a teacher could provide students as they complete their solutions. Leidl et al. used Google Analytics to collect information on users' activities in ScratchJr, a tablet-based BBPE that adapts the basic ideas of Scratch for a younger audience of ages 5-7 [29]. The work documents users' behaviors in a quantitative manner, deriving insights such as the average time spent by a user in ScratchJr and the types of programming blocks that are used more frequently by users [44]. To allow for richer interpretations of students' actions and to aid assessment of how they computationally think and problem solve, Grover et al. [33] also proposed a framework that uses hypothesis-driven analysis to support BBPE learning analytics.

To our knowledge, no TA systems exist specific to supporting teachers using Scratch. Furthermore, support is lacking for teachers to monitor students' learning progress when using Scratch to teach programming. Recent research on studying and predicting students' proficiency as they learn to program in Scratch has been predominantly framed around code artifact analysis [8, 53, 65, 70]. Unfortunately, this approach does not provide details regarding how students arrived at their final code solutions. Despite the increased interest in efficiently collecting data on students' learning process through integrated development environments in CEd research [39], such efforts involving Scratch have been ultimately

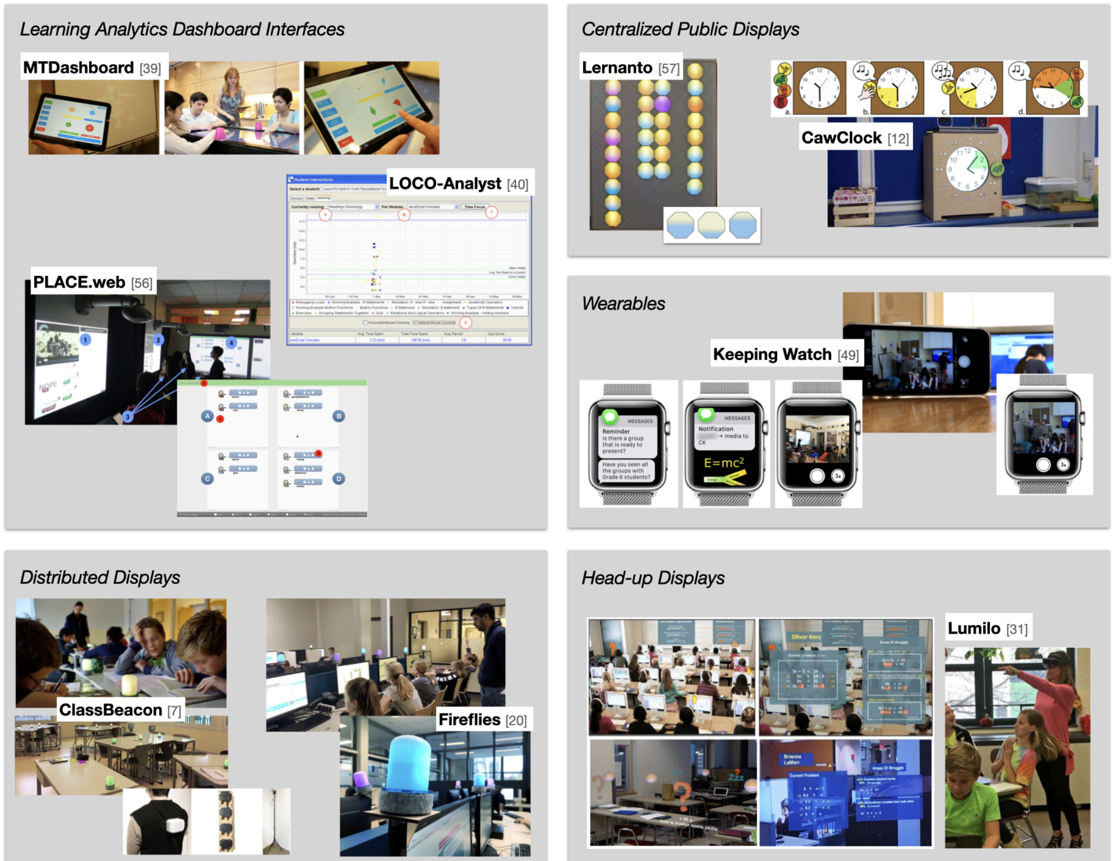

Figure 1: Examples of forms TA systems have taken in recent works.

time-consuming and non-scalable, with proposed methodologies involving mediums such as large file size screen recordings and Adobe Flash Player (no longer supported) [28, 47].

To address the research gap in monitoring how students learn programming with Scratch in a scalable way, Kong and Pollock introduced an automatic logging system to collect data on the steps that students take to develop their final code artifacts [41]. They also introduced a semi-automatic mining workflow that extracts clusters of programming behaviors that students commonly exhibited in the context of their data set. Such scalable logging and mining methodologies can be leveraged in the design of a Scratch TA system that helps with monitoring students' learning progress while supporting teachers' views on TA systems. However, there lacks a user-centered investigation on how such mechanism could be introduced to stakeholders such as teachers, which we aim to address with this work.

3 TEACHING AUGMENTATION FRAMEWORK

Our research on TA systems for teachers teaching Scratch programming is guided by the TA framework proposed by An et al., which provides system designers with a shared lens through which various designs can be described and analyzed [13]. The framework consists of five dimensions, each describing crucial design decisions that have not received significant attention in prior work: *augmentation target*, *attention*, *social visibility*, *presence over time*, and *interpretation*.

Augmentation target describes the teaching abilities that a TA system augments. On one end of the spectrum is *Perception*, which augments the teachers' situational awareness of their classrooms. On the opposite end is *Action*, which augments the set of actions that teachers choose to take given a situation, including automation of some of the actions that teachers would take. Along the spectrum is teachers' prioritization of classroom situations to act upon while teaching. In this dimension, there exists a tension between teachers' needs for autonomy versus automation. The framework advises designers to take time to understand the roles that teachers prefer to take in educational contexts and how automation could help or hurt while ensuring that their systems balance autonomy and automation to benefit teachers and students [34, 37].

Attention specifies which attention level(s) a TA system targets based on the concept of *attention continuum* [15]. TA systems could be designed to require that teachers provide little focal attention (*towards the periphery*) or continuous focal attention (*towards the *center of attention*). For this dimension, the TA framework notes that a teacher's attention can be scarce during an ongoing class session [35]. Thus, designers are recommended to effectively blend peripheral and focal display modes to ensure that TA systems can be accessible to teachers at multiple levels of attention.

Social visibility specifies how much information should be visible to each of the classroom stakeholders involved in the use of a TA system. A system could be designed to only make information visible to the teachers and thus have a *low social visibility*. In contrast, the system may allow everyone in a class, including students, to access its information, achieving a *high social visibility*. A TA system that aims to achieve a balance between low and high social visibility would allow teachers to control which information is shared with which stakeholders [35, 37, 49]. TA system designers should expect boundaries between teachers and students regarding the social visibility of any real-time information about classroom activities. The framework recommends that designers investigate form factors that support a sense of shared awareness to foster student motivation and collaboration [37].

Presence over time indicates how continuously a TA system should be present during an ongoing class session. Specifically, this dimension determines whether insights provided via augmentation should be *sometimes present* or *always present*. TA systems can be designed to only provide specific information either upon a teacher's request or in certain situations. Alternatively, a TA system can continuously provide information without the need for the teacher to intentionally seek and access it. A key challenge for TA system designers is allowing the system to help teachers in an opportunistic way, while ensuring such opportunism does not divert teachers away from the classroom.

Finally, the ***Interpretation*** dimension poses the question of how the task of interpreting phenomena in a classroom space should be divided between the teacher and the TA system. The system can provide less-processed information in a way that teachers have more freedom for interpretation. Alternatively, information could be pre-interpreted by the system, making it more immediately useful but leaving less room for interpretation by teachers. To address this tension, designers are encouraged to find an effective "blend" of human and machine intelligence [34].

4 STUDY 1: SEMI-STRUCTURED INTERVIEWS

To better understand how K-8 teachers utilize Scratch in their teaching practices and identify their preferences for a TA system, we first conducted a semi-structured interview study with ten K-8 teachers, which was approved by our Institutional Review Board (IRB). Detailed interview protocols are available in our supplementary materials.

4.1 Methodology

Our semi-structured interviews were conducted online. Participants were recruited via criterion sampling [57], resulting in a sample of teachers with at least one school year of experience teaching programming with Scratch in formal class settings (i.e., during the school day). Participating in our interviews were 10 K-8 teachers across different school districts in the eastern United States. Half of our interviewees (5/10) were teaching in the state of Delaware at the time of the interview, while the remaining 5 were teaching in Maryland. Our participants ranged in years of overall teaching experience ($M = 19.2; SD = 9.33$), and all 10 teachers are current or former partners with our or neighboring institutions' professional development programs. Table 1 describes the 10 teachers' demographics in greater detail.

Each interview involved two sets of questions, lasting an average of approximately 56 minutes ($SD = 16.86$), and was held by teleconference upon each teacher's consent. In the first set of questions, teachers were asked about their teaching experience and use of Scratch, including their levels of comfort with using Scratch and how they integrate Scratch into their class. Teachers were also asked questions on how they monitor students' progress when using Scratch and what teaching role they typically see themselves taking when teaching with Scratch (e.g., lecturer or facilitator). This first portion of the interview study was designed with the intention of understanding the various Scratch-involving teaching experience the 10 teachers brought to our interview space.

In the second part of the interview, teachers were asked questions about their preferences in a potential Scratch TA system design, based on the TA framework described in Section 3 [13]. We first described potential designs for Scratch TA systems, introducing the concept of teaching augmentation and showing them images of various forms that TA systems have taken in recent years, such as dashboard interfaces (e.g., [49, 50, 71]), wearables and head-up displays (e.g., [38, 63]), distributed displays (e.g., [11, 25]), and centralized public displays (e.g., [16, 72]). Then, the teachers were asked to describe the Scratch-involving teaching tasks in which they saw the greatest need for augmentation, preferred form factors for such systems, how data should be presented, and to what extent the task of interpretation should be divided between teacher and system.

To analyze the collected interview data, we first transcribed 9.3 hours total of video and audio-recorded interviews. Interview transcription involved a mix of naturalized and denaturalized techniques; transcriptions included pauses in speech, but not all forms of speech interventions were included (e.g., noises, accents) [55]. Transcripts were then coded for themes of interest using a mix of inductive and deductive codes [27]. The code manual was informed by our study protocol and literature such as the TA framework [13]. Initial codes included teachers' levels of comfort with Scratch (based on a 5-point Likert Scale, teachers' use of Scratch features in classrooms (e.g., studio, remix, comments, tutorials), and their need for augmentation (e.g., student assessment, emotional support, engagement). Our coding process also involved a *Gold Standard/Master Coder* approach [69]. The first author took the role of a *master coder* and a second member of the research team served as the *reliability coder*. We randomly selected 2 of 10 transcripts to code together. We used Krippendorff's alpha (α) to measure inter-rater reliability as it is applicable to a great diversity of data [42], with the unit of analysis being an answer or stream of answers for a certain topic.

With the initial version of the code book, we achieved an $\alpha = 0.68$ ($SE = 0.06$). Krippendorff suggests that although an $\alpha \geq .800$ is used as a threshold of good reliability, an $\alpha \geq .667$ can be used when tentative conclusions are acceptable. The two coders met to resolve disagreements and update the code book before independently coding the next interview. This round achieved an $\alpha = 0.70$

Table 1: Teacher Interviewees Demographics

Teacher	Gender(s)	Race	Experience (years)	School Type	Grade Level	Subject(s)
1	Cis Woman	White	22	Public	Elementary	Technology
2	Cis Woman	White	2	Public	Middle School	Computer Science
3	Cis Woman	White	27	Public	Elementary	Technology
4	Cis Woman	White	24	Public	Middle School	Business
5	Cis Man	White	26	Public	Elementary	Computer Science; English Language Arts
6	Cis Man	White	19	Private	Elementary	Technology
7	Cis Woman	White	24	Public	Elementary	Technology; STEM
8	Cis Woman	Black	8	Charter	Elementary	Technology
9	Cis Woman	White	30	Public	Elementary	Elementary across subjects
10	Cis Woman	White	10	Public	Middle School	STEM

($SE = 0.06$). Disagreements were again resolved, the code book was updated, and the master coder coded the remaining interviews.

4.2 Results

We present the emerging themes found in the interviewed K-8 teachers' responses, first regarding how they use and have used Scratch in their teaching practices, then the preferences they expressed for a potential Scratch TA system.

4.2.1 Part 1: Scratch Teaching Practices. We first summarize five main themes found in teachers' experiences with using Scratch both inside and outside of their classrooms.

Levels of Comfort with Scratch. On a 5-point Likert Scale (very uncomfortable to very comfortable), five teachers were comfortable with Scratch, a common reason being that there were students who surpass expectations and come into the classroom with more experience with Scratch and that there is always more that a teacher can learn. For instance, Teacher 1 (T1) stated, *"I'm comfortable enough to do basic projects. I'm a good critical thinker, so I can figure most out. That said, there are students that have far surpassed my ability because they've been in camps and classes for years."* Four of the remaining teachers expressed that they are very comfortable with using Scratch in their teaching practices, while one teacher felt neither uncomfortable or comfortable.

Using Scratch in Classrooms. All ten teachers expressed that they fuse Scratch-involving activities with other subjects, such as Literacy, Science, and Mathematics. Four teachers revolve students' Scratch experiences around a capstone project that students are expected to complete at the end of their Scratch unit. Specifically, the students complete various Scratch activities covering concepts that would be beneficial to them toward the end of their Scratch unit, when they would be expected to culminate those covered concepts into a larger Scratch project. Eight teachers implement subject-integrated Scratch activities at smaller scales, with some reporting to have adapted publicly available Scratch-based curricula to do so, such as *CS First* [2], *Creative Computing* [14], and *Scratch Encore* [30]. Some (3) also explained that they have adapted materials provided by their colleagues or school districts.

Monitoring Students' Scratch Progress. The interviewed teachers' ways of monitoring students' progress when learning programming using Scratch varied. Six teachers stated that they walk around the class during in-class Scratch activities to answer any questions students may have and/or examine how students are working on the activities. Four teachers also make use of remote monitoring and management systems, such as *NetSupport* [4], to observe student performance. Four teachers ask their students to upload their Scratch projects to designated studios as a way of monitoring their progress with projects and/or levels of understanding. Teacher 7 (T7), for instance, expressed that students are asked to upload the Scratch projects before they are at a state of completion and check each student's project throughout a Scratch lesson to see how much they have worked on the project. T2 expressed that they use studios slightly differently; studios served as a way of tracking how well students are understanding concepts covered in Scratch lessons.

Pros of Scratch. There were several benefits of Scratch in classrooms that were mentioned by the teachers during their interviews, such as its tutorials (T7) and its user-friendliness (T8). T6 also stated that being able to access Scratch for free outweigh some of the limitations that the teacher believed Scratch had in a classroom context. In addition, T4 has made use of Scratch in other languages in their classrooms (e.g., Spanish) and has found that translations of Scratch—Scratch is available in over 50 languages —helped address language barriers. As T4 explained, *"What I love about Scratch, too—if you have students that have difficulty reading or students that speak another language, it's so easy to use, even if you have those roadblocks in the way."*

Cons of Scratch. Overall, teachers predominantly saw the need for improvement in making Scratch friendlier for use in classrooms. For instance, Scratch's push for collaborative interactions with other Scratch users raises challenges in creating a safe digital learning environment for students. As a result, this led to a split in terms of those teachers who enable collaboration features (e.g., remix, comments) (7) and those who do not (5).

4.2.2 Part 2: Preferences in a Scratch TA System. We describe the interviewed teachers' expressed thoughts on a Scratch TA system using the five main dimensions of the TA framework [13], and summarize with overall benefits and issues the teachers see with Scratch TA systems.

Target. To determine which Scratch-specific teaching abilities should be augmented by a Scratch TA system, we asked the interviewed K-8 teachers two main questions, with the aim of utilizing teachers' responses to find healthy balances between *autonomy* and *automation* of teaching tasks. Here, we elaborate on these questions and teachers' responses.

During the interview, teachers were asked to describe Scratch teaching tasks for which they saw the greatest need for such augmentation, as teachers' need for autonomy may vary depending on the types of tasks [37]. Six participants in our study stated that they would like TA systems to help achieve high engagement levels among students as they work on Scratch-involved activities (3) and maintain and/or enhance students' motivation to learn programming via Scratch (5). T2, for instance, stated: *"student motivation, knowing who was more motivated or, 'he was getting stuck on a certain area,' because not all the students will tell you if they're kind of stuck on something right away. So definitely some sort of monitor for that."* Five teachers also stated that it would be helpful for Scratch TA systems to augment teachers' abilities to provide emotional support, with T1 and T10 specifying that frustration among students was an emotion they wished to address.

Four teachers wanted future Scratch TA systems to augment the task of assessing students' work and/or learning progress. Teachers' ideas for such assessment-related augmentation came in various forms. For instance, T10 expressed that they *"really do struggle with coming up with good assessments and rubrics,"* and indicated that Scratch TA systems could guide how to best assess their students.

Some participants (4/10) also expressed that they would like for Scratch TA systems to help both teachers (3) and students (2) with effective time management during class, whether it be for teachers to best spread their time among their students when providing in-class support or for students to manage their time while working on their Scratch projects. Other participants (2/10) also specifically saw a need for teaching augmentation when getting *all* students started with using Scratch (e.g., to account for differences in levels of understanding, to monitor whether students have properly logged into their Scratch accounts for class).

Teachers were also asked during the interview about the roles they take in the class when teaching using Scratch. Such insight is important in determining how to effectively balance teachers' autonomy with automation for future Scratch TA system designs, as teachers' needs for autonomy may differ between those who see themselves predominantly working in lecture-heavy contexts and those who position themselves as facilitators [12, 25]. Six teachers saw themselves taking both the roles of a lecturer and a facilitator, providing students with concrete instructions as needed and also allowing students to work at their own pace. Four participants strictly described themselves as more of a facilitator in lessons that are self-paced. Specifically, both T3 and T5 encourage students to take leadership roles in their classrooms, allowing students who are very comfortable with Scratch to help peers who are not or would normally shy away from asking teachers for help.

Attention. To explore what levels of attention the interviewed teachers would be comfortable with sparing with a Scratch TA system, we asked participants what their preferred and unfavorable forms of augmentation would be. Participants took the most interest in distributed displays such as ambient digital lamps (7/10) and half the group also expressed that augmentation could come in the form of a dashboard interface (5/10). No participants mentioned that they preferred that augmentation is available as a centralized public display. There were also teachers who were interested in wearable devices (3/10), with T4 finding it beneficial to be able to access augmentation anywhere in their classroom using wearables such as smartwatches. In addition, two teachers were captivated by the idea of augmentation being available in the form of head-up displays.

At the time of the interviews, no teachers mentioned dashboard interfaces as an unfavorable Scratch TA system design. Three teachers were not in favor of using wearables in their classrooms, with T1 stating that they may find it distracting while teaching, and T3 and T8, who are currently smartwatch users in their daily lives, not being sure about using them for teaching purposes. There were also teachers who did not see themselves using head-up displays in their classrooms (3/10). Some were uncertain about TA systems that take the form of centralized public displays (3/10) and distributed displays (2/10) because they could be costly or anxiety-inducing.

Social Visibility. When asked with which stakeholders augmented awareness of a classroom should be shared, all teachers agreed on allowing information from Scratch TA systems to be visible to students. Four teachers expressed that they would be okay with students having full access to the types of information, but two specified that each student should only be able to see information that is specific to them, as seeing that of peers or knowing that their information is being shared with their classmates may be a form of disrespect to their privacy. Several teachers, on the other hand, believed that students should have access to some, but not all, types of information offered by a TA system (3). T9, for instance, stated that they would show students a system's information when providing feedback (*"I would use the data in either a one-on-one conference or a small group conference to show growth and just set growth."*). Two preferred future TA systems not come with a default, but rather allow teachers to set levels of access and visibility to students.

Presence over Time. To achieve a sound balance in how continuously augmentation is available to teachers, participants were asked whether they thought augmentation should always be present (e.g., continuously available for intentional or opportunistic access) or sometimes present (e.g., available on-request or contextually). Five teachers voiced that the presence of augmentation should be customizable according to what individual teachers need within the context of their classrooms and teaching practices. Two teachers favored that augmentation always be present in their classrooms (e.g., T10 stated, *"I think it's a tool that's going to help us and if it's a tool that helps, I like the idea of being able to know where my kids are, and that's something I'm going to always want to know."*). Conversely, T6 wanted augmentation to be sometimes present (*"Sometimes. Not all the time. It could become a distraction."*). In addition, two other teachers were unsure, with T4 stating that they would not know until they tried utilizing TA systems.

Interpretation. Interviewed teachers were asked to describe their thoughts on the extent to which the task of interpreting students'

Scratch activities should be divided between a teacher and the system. Half the group of interviewed teachers specified that they would want to take on more of the interpretation task, one of the main reasons being that information provided by the system has to be interpreted according to the varying learning needs of their students. T7 elaborates on this further:

> *I definitely think [interpretation] needs to be differentiated per child, whether you have somebody who's [in] special needs or ESL [English as a second language]. That all needs to be taken into factor and I don't know if a computer can automate that. You might have a student who has mental health issues who needs extra support in that area. So, I think the empathy part of the teacher needs to be involved in that. (T7)*

Some teachers thought that the interpretation task should be a collaborative effort between the teacher and the system, where a teacher is able to make their own judgements about information as a result of augmentation, but still be given some automated suggestions that could give teachers a starting point for interpretation (3). For instance, T5 mentioned that they would perceive the systems' interpretation as an *"unbiased view of things."* T4 elaborated, (*"I think every teacher has bias, and I think [a TA system's interpretation] is a non-biased way to look at the data"*).

Benefits of Augmentation. To conclude the interview, the teachers described various ways in which they believed a Scratch TA system may benefit both teachers and students, if any. Mentioned benefits included being able to do the following: monitor students' progress when learning programming with Scratch (3); adapt their lessons according to students' exhibited learning needs (3), and; provide better-quality support (2) and feedback (2) to students before they feel frustrated (T10), for instance, by dividing focus appropriately (T9). One teacher, T5, also saw the potential for such systems, if designed according to both teachers' and students' needs, to encourage more equitable teaching practices.

Issues of Augmentation. Teachers also listed potential issues that an introduction of a Scratch TA system may bring to a classroom. For instance, students may have an increased sense of competition against their peers, especially if the awareness of a class' activity is shared among both teachers and students (3). Cost of TA systems, especially when they are physically present in the classrooms, may serve as a barrier against teachers utilizing the systems to augment their teaching practices (3). The presence of TA systems may be a distraction to some students (T4) and also invoke emotions among students that may be harmful to their learning experiences (4), such as discomfort (T2), stress (T5), and anxiety (T1). Some stakeholders may also view Scratch TA systems' use of information of their learning activities as a privacy invasion (2). In addition, students may attempt to win the TA system and ultimately not fully benefit from learning with Scratch (T3), and both students (T5) and teachers (T10) may become dependent.

5 STUDY 2: SURVEY

To contextualize our interview findings with a broader audience, we conducted an anonymous online survey using Qualtrics. The survey study was also approved by our IRB.

5.1 Methodology

Using snowball sampling, we recruited teachers who have taught using Scratch primarily at the K-8 level in at least one formal class session in the United States. To distribute the survey, we first directly invited K-8 teachers in our 2022 cohort of our institution's annual CS professional development program. We then welcomed the teachers to share the survey with their colleagues who met our criteria.

The survey was designed to serve as a complement to the qualitative data from the interviews. Specifically, each question was composed around the key themes that emerged from the interviewees' questions. First, the teachers were asked to watch a brief video that described the purpose of the survey, as well as what is implied by "teaching augmentation." The teachers were then asked about their preferences in regards to a Scratch TA system and the various forms a TA system can potentially take. The survey concluded with demographic questions to understand our respondents' background, teaching experience, and school environment. Each question offered closed-ended response options based on the interview themes, as well as open-ended options that allowed survey participants to elaborate on their responses. Survey questions are provided in the supplementary materials.

Participants were compensated for their participation using an incentive scheme. The first 50 survey respondents who met our criteria and completed all of the survey questions were eligible to receive a $50 Amazon gift card. Respondents were also given an additional $10 Amazon gift card for every person they recommended to take the survey provided that the referred respondent(s) also fully completed the survey. Each respondent could receive additional gift cards for up to 5 referred teachers (i.e., maximum $100 total).

5.2 Data & Analysis

After removing responses from bots using fraud detection measures (see [61]) and incomplete responses, 37 respondents met our inclusion criteria and finished the survey. None were repeat participants from the interview study.

5.3 Participants

Our pool of respondents was made up of 21 cis women and 16 cis men, 35 of whom identify as not having a disability or other chronic condition (2 preferred not to disclose). The participants closely identify as White (29), Black or African American (9), American Indian or Alaska Native (3), Asian or Asian American (3), Hispanic, Latino, Latina, or Latinx (3), Middle Eastern or Northern African (2), and/or Native Hawaiian or other Pacific Islander (1). The survey respondents brought with them an average of 10.27 years of overall teaching experience (SD = 8.79). Their prior experience teaching using Scratch in formal class settings at the time of the study varied; 11 teachers have used Scratch for 6 to 12 months, 9 teachers for 1 to 2 years, 7 teachers for 2 to 5 years, 5 teachers for more than 5 years, and 5 teachers for 6 months or less.

All 37 respondents teach or have taught using Scratch primarily at the K-8 level (i.e., 9 at grades K-2, 15 at grades 3-5, 18 at grades 6-8) and, according to the United States Census Bureau [22], represent 20 Southern, 10 West, five Midwest, and two Northeast states. They teach various subjects in various types of schools—21 teach at a public school, 15 at private schools, and one at a charter school—in suburban (23), urban (13), and rural (1) settings. The majority (30) of respondents' schools are predominantly made up of White students, while four are made up of Black or African American students and two American Indian or Alaska Native (one teacher responded that they were unsure). More than half of respondents (19) teach students who have an Individualized Education Plan or a 504 Plan due to disabilities or special needs, and 18 teachers teach students identified as English-language learners (ELL), also known as limited-English proficiency.

5.4 Results

Here we describe key findings from our survey, first according to the TA framework dimensions, then report on additional comments from the respondents on the impacts of augmentation on BBPE instruction and learning.

5.4.1 Augmentation Target. On a 4-point Likert scale from "Not at all" to "Extremely" useful, the respondents indicated how useful they would find a TA system should it be designed to meet each of the augmentation needs expressed by the interviewed teachers from the previous study. All nine types of augmented needs presented in the survey were rated as either "Extremely" or "Somewhat" helpful by a majority of the respondents (32/37; 86%). The most favorably rated augmented needs include: providing better feedback (e.g., informing a student what they have done well when programming in Scratch and what skills they could improve upon); maintaining and supporting students' engagement when learning Scratch; monitoring students' progress when learning Scratch; and adapting Scratch lessons according to students' learning needs. Two teachers described an additional teaching task specific to Scratch that they would like to have augmented support for, which is to be able to better monitor how and where students dedicate their attention to within the Scratch website (e.g., *"monitor the percentage of time that students spend working in their own projects, as opposed to playing other people's Scratch games from the gallery"*). Examples of augmentation needs that were deemed not as helpful by the teachers were assessing students' Scratch programming skills, as well as managing both the teachers' and students' time while teaching and learning to program using Scratch in class respectively.

5.4.2 Social Visibility. When asked with which stakeholders augmented awareness of a classroom should be shared, the respondents varied in how much information that they believe individuals and the class should be able to access. Approximately half of the respondents (19/37; 51%) stated that as stakeholders, students should have *full* access to information on their *individual* Scratch performance shown in a TA system (e.g., amount of time a student spends on a Scratch project, automated predictions of a student's Scratch programming strengths/weaknesses). Less than half of respondents (16/37; 43%) expressed that students should only be able to access *some* types of such information offered by a TA system while only two participants specified that visibility of such information should be customizable by teachers. The respondents' thoughts on levels of access and visibility to students differed in whether a TA system should only offer aggregate information of the *whole class'* Scratch performance shown in a TA system (e.g., a student can observe their class' average Scratch performance or a student should also be able to observe each classmate's Scratch programming learning pace). More than half of the respondents (21/37; 57%) thought students should only have some access to such information, 22% believed that students should have full access (8/37), 16% stated that TA systems should allow teachers to customize students' access to such information (6/37), and the remaining 5% expressed that students should have no access at all (2/37).

5.4.3 Presence over Time. When asked about their preferences regarding the extent to which augmentation should be continuously available to teachers over time, the majority of the respondents (26/37; 70%) preferred that teachers be able to customize an augmentation's presence. A respondent further clarified that *"teaching at any level is a constant state of recognizing and observing student needs in real-time and adapting your style and level of instruction accordingly,"* and thus being able to customize would be of utmost importance. Another added that a TA system's presence *"should be customizable so that a teacher may choose when to use it, but at the same time, feel that [the] TA [system] should always be (readily) available for an educator to use at any time."* On the other hand, 16% (6/37) stated that they would want augmentation to sometimes be present (e.g., available on-request or contextually) while 14% (5/37) preferred for it to always be present (e.g., continuously available for intentional or opportunistic access).

5.4.4 Interpretation. The survey respondents' thoughts around interpretation tasks differed from those of the interview study participants, out of whom half the group expressed that they would want to take on more of the task. In the survey, being able to tailor TA system features to each teacher's unique needs was a theme observed in approximately half of the respondents' preferences regarding a TA system's involvement in the task of interpreting students' Scratch activities (20/37; 54%). The remaining respondents were split on whether a teacher or a TA system should take on more of the interpretation task; 27% (10/37) wanted teachers to take on more of the task while 19% (7/37) preferred that the TA system do more. One of the respondents elaborated on their preference for a TA system to take more charge of the interpretation task, *"In an ideal world, the teacher should have more of a role in interpreting. The reality of teaching, however, is that teachers need quick and efficient tools so that they don't have to spend a lot of time interpreting, and quickly read and understand data."*

5.4.5 Attention. Survey respondents indicated to which extent they preferred various forms of TA systems using a 5-point Likert scale ("Never" to "A great deal"), then had the option to elaborate on their ratings for each of the TA forms. The respondents' ratings are described in greater detail in Table 2. Overall, the respondents favored dashboard interfaces the most. Some of the reasons for their preference for dashboard interfaces included being able to more easily monitor student progress and then accommodate students better accordingly. A respondent specifically brought these points up in the context of scalability (i.e., a dashboard makes personalizing instruction with a large group of students possible in a time-efficient manner). Another respondent mentioned their experience with

Table 2: Survey respondents' preferred forms of TA systems according to a 5-point Likert scale ("Never" to "A great deal").

	Ratings from survey respondents (n=37)									
	A great deal		Much		Somewhat		Little		Never	
TA system forms	N	%	N	%	N	%	N	%	N	%
Dashboard Interfaces[a]	9	25%	19	53%	5	14%	2	6%	1	3%
Distributed Systems	10	27%	12	32%	8	22%	5	14%	2	5%
Smartwatches	5	14%	8	22%	17	46%	5	14%	2	5%
Head-up Displays	10	27%	12	32%	8	22%	4	11%	3	8%
Public Peripheral Displays	9	24%	11	30%	12	32%	3	8%	2	5%

[a]We only report on n=36 ratings for dashboard interfaces as 1 response was missing a rating for this TA form.

using the analytics feature on CodeMonkey [7], an online game-based learning environment, which they felt was *"good to see the progress the [students] have made and the areas where they are having difficulty."* They then noted their curiosity as to how Scratch-specific analytics would work unless there was a set of developmental skills defined, especially given its open-ended nature.

The respondents were split on their preferences for distributed systems, head-up displays, and public peripheral displays. The respondents listed unique benefits for each of the three TA forms. Specifically, they felt that: distributed systems could serve as an alternative communication medium for students (i.e., ELL, introverted, or shy students could ask for help without needing to verbalize the request); head-up displays would not require teachers to divert attention away from students while accessing data on student progress offered by a TA system; and public peripheral displays could increase engagement of some students. The respondents, however, specified that the three TA forms all share the same drawback, which is that they could be distracting to teachers and/or students and *"limit my ability to take advantage of 'teachable moments.'"* In addition, some respondents were concerned about making information on students' Scratch performance visible to their peers via forms like distributed systems and public peripheral displays.

TA systems in the form of a smartwatch were favored the least by the respondents. Although the respondents thought that smartwatches could create useful opportunities to assist and support novice teachers and those not comfortable teaching certain concepts, smartwatches were also viewed as distracting to both teachers and students. Additionally, they might prevent teachers from delivering *"direct and quick intervention"* at their fullest potential and take away meaningful discussion opportunities, for instance, when teachers deviate their attention from the students and classroom to a smartwatch.

5.4.6 Benefits and Issues of Augmentation. To consider how future BBPE-based TA system designs impact teaching and learning experiences, the respondents were asked to describe how they see their preferred form of TA system impacting learning goals and/or outcomes. The respondents predominantly held positive views toward the systems' potential impacts in classrooms. Some of the mentioned positive traits of the respondents' preferred TA system designs included having *"a quicker and better understanding [as teachers] of student performance or mastery of skills,"* being able to *"enhance students' interest in learning"* and *"implement personalized"* and *"targeted teaching,"* as well as enabling *"better communication and dialogue with students."*

The respondents also pointed out challenges and barriers that may come with the implementation of TA systems in BBPE instruction. Logistically, the respondents were concerned about the funding for and cost of the TA systems, as well as the additional time and training that may be required of teachers in order to properly adopt TA systems. Some respondents mentioned various stakeholders' reactions to the integration of augmentation (e.g., school administrators and parents' impressions of TA, increase in distraction among students) as well. However, when asked about how likely their school administrators would allow the respondents to use BBPE-based TA tool(s) in their classrooms, using a 5-point Likert question ("Extremely unlikely" to "Extremely likely"), the majority of respondents answered either "Extremely likely" (49%) or "Somewhat likely" (46%), with only 5% answering "Neither likely nor unlikely."

6 DISCUSSION

Both interviewed and surveyed teachers discussed several Scratch teaching tasks that they thought a TA system might help with. Both groups of teachers favored TA systems that are able to help maintain and enhance students' levels of motivation and engagement, and adapt Scratch-involving lesson plans according to students' expressed learning needs. The interviewees specifically exhibited preference for autonomy during such Scratch teaching tasks, given that all 10 take the role of a facilitator during Scratch-involving classroom activities.

We found that our participants' expressed augmentation needs were in alignment with empirical findings on common pedagogical needs among K-12 computing teachers, one of which is to receive guidance on implementing student-centered teaching practices [67]. By introducing BBPE-based TA systems designed to meet augmentation needs to achieve student-centered teaching, students may feel more confident and positive about their BBPE learning experiences. Further, TA systems designed to support student-centered teaching practices may also help reduce student anxiety and fears of failure. Such a system would thus support our research's overall motivation to assist teachers in creating a greater sense of computing self-efficacy among their students when learning about the domain through block-based programming. For example, TA systems in the form of a dashboard interfaces could display to teachers easy-to-grasp summaries of information on students' motivation and engagement (e.g., measured through their activities and time spent within a BBPE as described by some survey respondents). Should teachers make use of such information to more productively divide their time, attention, and support among students, especially for those who are at risk of or have developed a sense of anxiety and/or a fear of failure while learning with a BBPE, the overall levels of motivation and engagement—and ultimately students' self-efficacy toward computing—may improve. Using BBPE-based TA systems in this manner, teachers could potentially receive more support to achieve a healthy balance between direct instruction and open-ended exploratory learning in their pedagogy.

In both interview and survey studies, the participants' most preferred augmentation form was a dashboard interface. The remaining four examples of TA forms (distributed displays, head-up displays, wearables, centralized public displays) received mixed reactions across the two studies, with each augmentation form bringing in benefits that are unique to its traits and helping to meet some of the teacher participants' expressed needs, but ultimately carrying various concerns. The participants' concern that augmentation could end up being distracting to teachers and/or students should it come in any of the four forms was especially in contradiction to observations from prior work on augmentation outside of the BBPE context. For example, pilot and user exploratory studies conducted with previously introduced TA systems in some of the four forms revealed that teacher participants did not consider the TA systems to be a distraction to the teachers and students (e.g., [11, 16, 63, 72]). The use of several existing TA systems has been with what some called 'first-time enthusiasm,' where there exists high levels of enthusiasm with the introduction of a new technology, but the enthusiasm decreases as it becomes more part of teachers' and/or students' routines [16].

All participants in both interviews and survey responses believed that students, as stakeholders, should also have access to information offered by a TA system to some extent. However, teachers' opinions on students' levels of access varied. Some were open to the idea of students having full access while others felt that students' levels of access should be controlled. Their views also differed based on whether a TA system was providing information on each individual student or the whole class' Scratch performance.

The two studies' participants had mixed preferences on to what extent the task of interpreting in-class phenomena on students' BBPE learning should be divided between a teacher and a TA system. 54% of the survey respondents wanted to share the task with the TA system, while half of the interviewees favored taking on the task more.

In addition, the teachers felt that there were several benefits and issues to the introduction of a TA system. They see potential for augmentation to benefit both teachers and students should a TA system be successfully designed to meet both stakeholder groups' needs, but also see TA systems' potential for interrupting students who are learning programming using Scratch. While these findings around teachers' preferences for Scratch-based TA systems may not be representative of the greater teaching community, they suggest that teachers would find TA systems beneficial if their expressed teaching needs are met.

6.1 Design Recommendations and Future Work

Based on the results from our studies focused on teaching programming with Scratch and using the TA framework by An et al. [13], we make design recommendations for TA systems, which are broadly applicable to BBPEs like Scratch.

6.1.1 Augmentation Target. For teachers like all 10 of our interviewees, who act more as facilitators when conducting Scratch-involving lessons, with some occasionally mixing in lectures, we recommend that TA systems be designed for their more likely use of helping to extend a teacher's ability to monitor happenings in their classrooms. Teachers gave this monitoring capability a higher priority than more than receiving automated suggestions from TA systems on what types of actions to take. However, the interview and survey participants also provided insight into other teaching tasks that they would like for a BBPE TA system to assist them with including: engagement, motivation, and implementing student-centered practices. TA system designs should also consider how to help with these important tasks.

6.1.2 Attention and Presence over Time. Based on how beneficial or unfavorable the interview participants viewed the various forms of TA systems, it may be most beneficial for K-8 school settings if BBPE TA system designers focus on providing teachers augmentation in the forms of dashboard interfaces and distributed displays in future work. A combination of both forms is also recommended to make augmentation available at multiple attention levels.

6.1.3 Social Visibility. Future TA systems should support views for both teachers and students. Overall, our participants believed that augmented awareness of a classroom should be shared with students, who are also one of the main stakeholders of TA systems. However, there were split opinions on whether students should have some or full access to the information provided by a BBPE TA system. Follow-up discussions with a larger group of teachers and students in various class and school settings are needed to concretely determine what would be best for both stakeholders.

6.1.4 Interpretation. In the K-8 context, we believe TA system designs may benefit teachers the most if human interpretation is prioritized while information is still presented in a way that teachers would not have difficulty comprehending it. Several user and prototype-based studies may be needed to identify how information should best be presented to teachers. Given some teachers' perception of BBPE TA systems serving as an unbiased perspective, we also recommend future BBPE TA system designers to practice decentering themselves from dominant ideologies that contribute to the persistence of bias in society throughout the designing process, to prevent the risk of bias fusing into their designs and broadly, teachers' teaching approaches and students' learning experiences [24].

6.2 Limitations

There are several limitations of this work. First, the data that we present is U.S.-centric and may not reflect Scratch-involving teaching practices and preferences in other countries. Our semi-structured interview study involved a small sample of teachers, none from the same school, in two states in the United States, which should be taken into account when interpreting the results; the themes commonly found through our interviews are, by no means, exhaustive or definitive. We expanded our data through a survey as a quantitative supplement to the interview responses and intended for our survey to involve a large sample. Because we were recruiting a very specific demographic of participants—teachers who teach or have taught using Scratch at the K-8 level in the United States in one or more formal class sessions (e.g., during the school day)—we could not obtain a statistically significant set of responses. The teaching context of the participants may not represent all kinds of teaching environments; however, our samples did include public and private schools, and urban and rural schools. While our participant samples may not represent all teaching contexts, our study results indeed highlight potential design decisions and preferences of teachers for BBPE-based TA systems. Like any research study of this kind, we recommend replications of our study with larger audiences representative of other communities.

7 CONCLUSION

The potential of TA systems has been explored across several disciplines at the intersection of education and HCI. To ensure that augmentation has a positive impact on teaching practices, designs of TA systems should be guided by their stakeholders' expressed preferences. In this work, we conducted 10 semi-structured interviews and collected 37 anonymous survey responses from K-8 teachers to pinpoint their preferences in potential TA systems involving Scratch, a renowned BBPE. Based on our findings, we present design recommendations for future systems around the 5 key dimensions of the TA framework [13], such as best targeting teachers' attention level through their preferred TA forms (dashboards and ambient displays) and prioritizing teacher interpretation of students' behaviors while using BBPEs over that of an automated system. We encourage future TA system designers to consider our recommendations and how their designs can best assist teachers in meeting students' varying learning needs through data-driven instructional decisions. As a result, our work will be of interest to users of BBPEs including Scratch, teachers, school administrators, application designers, and researchers working on improving CEd learning environments.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1940700. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] 2019. *2019 State of Computer Science Education.* advocacy.code.org/
[2] 2021. CS First. csfirst.withgoogle.com.
[3] 2021. Lightspeed. www.lightspeedhq.com.
[4] 2021. NetSupport. www.netsupportschool.com.
[5] 2021. Scratch in Many Languages – Scratch Wiki. en.scratch-wiki.info/wiki/Scratch_in_Many_Languages.
[6] 2021. Securly. www.securly.com.
[7] 2022. CodeMonkey: Coding for Kids | Game-Based Programming. codemonkey.com
[8] Efthimia Aivaloglou and Felienne Hermans. 2016. How Kids Code and How We Know: An Exploratory Study on the Scratch Repository. In *Proc. of the 2016 ACM Conf. on Int. Comput. Educ. Res.* ACM.
[9] Efthimia Aivaloglou and Felienne Hermans. 2019. Early Programming Education and Career Orientation: The Effects of Gender, Self-Efficacy, Motivation and Stereotypes. In *Proc. of the 50th ACM SIGCSE Tech. Symp. on Comput. Sci. Educ.* ACM.
[10] Hamed Alavi and Pierre Dillenbourg. 2012. An Ambient Awareness Tool for Supporting Supervised Collaborative Problem Solving. *IEEE Trans. on Learn. Techn.* 5, 3 (2012).
[11] Pengcheng An, Saskia Bakker, Sara Ordanovski, Ruurd Taconis, and Berry Eggen. 2018. ClassBeacons: Designing Distributed Visualization of Teachers' Physical Proximity in the Classroom. In *Proc. of the 12th Int. Conf. on Tangible, Embed., and Embodied Interact.* ACM.
[12] Pengcheng An, Saskia Bakker, Sara Ordanovski, Ruurd Taconis, Chris LE Paffen, and Berry Eggen. 2019. Unobtrusively Enhancing Reflection-in-Action of Teachers through Spatially Distributed Ambient Information. In *Proc. of the 2019 CHI Conf. on Hum. Factors in Comput. Syst.*
[13] Pengcheng An, Kenneth Holstein, Bernice d'Anjou, Berry Eggen, and Saskia Bakker. 2020. The TA Framework: Designing Real-Time Teaching Augmentation for K-12 Classrooms. In *Proc. of the 2020 CHI Conf. on Hum. Factors in Comput. Syst.* ACM.
[14] Creative Computing Lab at the Harvard Graduate School of Education. 2021. Creative Computing Curriculum. scratched.gse.harvard.edu/guide/.
[15] Saskia Bakker and Karin Niemantsverdriet. 2016. The Interaction-Attention Continuum : Considering Various Levels of Human Attention in Interaction Design. *Int. J. Des.* 10, 2 (2016).
[16] Saskia Bakker, Elise van den Hoven, Berry Eggen, and Kees Overbeeke. 2012. Exploring Peripheral Interaction Design for Primary School Teachers. In *Proc. of the 6th Int. Conf. on Tangible, Embed., and Embodied Interact.* ACM.
[17] Albert Bandura, William H Freeman, and Richard Lightsey. 1999. Self-efficacy: The exercise of control.
[18] Ashok Basawapatna, Alexander Repenning, Kyu Han Koh, and Hilarie Nickerson. 2013. The Zones of Proximal Flow: Guiding Students through a Space of Computational Thinking Skills and Challenges. In *Proc. of the 9th Annu. Int. ACM Conf. on Int. Comput. Educ. Res.* ACM.
[19] David Bau, Jeff Gray, Caitlin Kelleher, Josh Sheldon, and Franklyn Turbak. 2017. Learnable Programming: Blocks and Beyond. *Commun. ACM* 60, 6 (2017).
[20] Matthew Berland, Taylor Martin, Tom Benton, Carmen Petrick Smith, and Don Davis. 2013. Using Learning Analytics to Understand the Learning Pathways of Novice Programmers. *J. the Learn. Sci.* 22, 4 (2013).
[21] Paulo Blikstein, Marcelo Worsley, Chris Piech, Mehran Sahami, Steven Cooper, and Daphne Koller. 2014. Programming Pluralism: Using Learning Analytics to Detect Patterns in the Learning of Computer Programming. *J. the Learn. Sci.* 23, 4 (2014).
[22] United States Census Bureau. 2013. Census Regions and Divisions of the United States.
[23] Steve Cooper, Shuchi Grover, Mark Guzdial, and Beth Simon. 2014. A Future for Computing Education Research. *Commun. ACM* 57, 11 (2014).
[24] Sasha Costanza-Chock. 2020. *Design Justice: Community-Led Practices to Build the Worlds We Need.* The MIT Press, Cambridge, MA, USA.
[25] Bernice d'Anjou, Saskia Bakker, Pengcheng An, and Tilde Bekker. 2019. How Peripheral Data Visualisation Systems Support Secondary School Teachers during VLE-Supported Lessons. In *Proc. of the 2019 on Des. Interact. Syst. Conf.*
[26] Emine Erdem. 2015. The Relationship between Self-Efficacy and Attitudes of Chemistry Teacher Candidates. *Probl. of Educ. in the 21st Century* 63, 1 (2015).
[27] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *Int. J. Qual. Methods* 5, 1 (2006).
[28] Daniel Amo Filvà, Marc Alier Forment, Francisco José García-Peñalvo, David Fonseca Escudero, and María José Casañ. 2019. Clickstream for Learning Analytics to Assess Students' Behavior with Scratch. *Future Gener. Comput. Syst.* 93 (2019).
[29] Louise Flannery, Brian Silverman, Elizabeth Kazakoff, Marina Umaschi Bers, Paula Bontá, and Mitchel Resnick. 2013. Designing ScratchJr: Support for Early Childhood Learning through Computer Programming. In *Proc. the 12th Int. Conf. on Interact. Des. and Children.* ACM.

[30] Diana Franklin, David Weintrop, Jennifer Palmer, Merijke Coenraad, Melissa Cobian, Kristan Beck, Andrew Rasmussen, Sue Krause, Max White, Marco Anaya, and Zachary Crenshaw. 2020. Scratch Encore: The Design and Pilot of a Culturally-Relevant Intermediate Scratch Curriculum. In *Proc. of the 51st ACM Tech. Symp. on Comput. Sci. Educ.* ACM.

[31] Neil Fraser. 2015. Ten Things We've Learned from Blockly. In *2015 IEEE Blocks and Beyond Workshop (Blocks and Beyond)*.

[32] Michail Giannakos, Spyros Doukakis, Ilias Pappas, Nikos Adamopoulos, and Panagiota Giannopoulou. 2015. Investigating Teachers' Confidence on Technological Pedagogical and Content Knowledge: An Initial Validation of TPACK Scales in K-12 Computing Education Context. *J. Comput. in Educ.* 2, 1 (2015).

[33] Shuchi Grover, Satabdi Basu, Marie Bienkowski, Michael Eagle, Nicholas Diana, and John Stamper. 2017. A Framework for Using Hypothesis-Driven Approaches to Support Data-Driven Learning Analytics in Measuring Computational Thinking in Block-Based Programming Environments. *ACM Trans. Comput. Educ.* 17, 3 (2017).

[34] Jeffrey Heer. 2019. Agency Plus Automation: Designing Artificial Intelligence into Interactive Systems. *Proc. of the Natl. Acad. Sci.* 116, 6 (2019).

[35] Kenneth Holstein, Gena Hong, Mera Tegene, Bruce McLaren, and Vincent Aleven. 2018. The Classroom as a Dashboard: Co-Designing Wearable Cognitive Augmentation for K-12 Teachers. In *Proc. of the 8th Int. Conf. on Learn. Anal. and Knowl.* ACM.

[36] Kenneth Holstein, Bruce McLaren, and Vincent Aleven. 2019. Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity. *J. Learn. Anal.* 6, 2 (2019).

[37] Kenneth Holstein, Bruce McLaren, and Vincent Aleven. 2019. Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms. In *Int. Conf. on Artif. Intell. in Educ.* Springer.

[38] Kenneth Holstein, Bruce M McLaren, and Vincent Aleven. 2018. Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms. In *Int. Conf. on Artif. Intell. in Educ.* Springer.

[39] C. Hundhausen, D. Olivares, and A. Carter. 2017. IDE-Based Learning Analytics for Computing Education: A Process Model, Critical Review, and Research Agenda. *ACM Trans. Comput. Educ.* 17, 3, Article 11 (2017).

[40] Dawn R. Johnson. 2011. Examining sense of belonging and campus racial diversity experiences among women of color in stem living-learning programs. *J. of Women and Minorities in Sci. and Eng.* 17, 3 (2011), 209–223.

[41] Minji Kong and Lori Pollock. 2020. Semi-Automatically Mining Students' Common Scratch Programming Behaviors. In *Koli Calling '20: Proc. of the 20th Koli Calling Int. Conf. on Comput. Educ. Res.* ACM.

[42] Klaus Krippendorff. 2018. *Content Analysis: An Introduction to Its Methodology.* Sage publications.

[43] Seanna Leath and Tabbye Chavous. 2018. Black Women's Experiences of Campus Racial Climate and Stigma at Predominantly White Institutions: Insights from a Comparative and Within-Group approach for STEM and Non-STEM Majors. *J. of Negro Educ.* 87 (2018).

[44] Kaitlyn Leidl, Marina Umaschi Bers, and Claudia Mihm. 2017. Programming with ScratchJr: A Review of the First Year of User Analytics. In *Conf. Proc. Int. Conf. on Comput. Think. Educ.* The Educat. Univ. Hong Kong.

[45] Rensis Likert. 1932. A technique for the measurement of attitudes. *Arch. of Psychol.* (1932).

[46] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to Program: Gender Differences and Interactive Effects of Students' Motivation, Goals, and Self-Efficacy on Performance. In *Proc. of the 2016 ACM Conf. on Int. Comput. Educ. Res.* ACM.

[47] Sze Yee Lye and Joyce Hwee Ling Koh. 2018. Case Studies of Elementary Children's Engagement in Computational Thinking Through Scratch Programming. In *Comput. Think. in the STEM Discipl.* Springer.

[48] John Maloney, Kylie Peppler, Yasmin Kafai, Mitchel Resnick, and Natalie Rusk. 2008. Programming by Choice: Urban Youth Learning Programming with Scratch. In *Proc. of the 39th SIGCSE Techn. Symp. on Comput. Sci. Educ.* ACM.

[49] Roberto Martinez-Maldonado, Judy Kay, Kalina Yacef, Marie-Theresa Edbauer, and Yannis Dimitriadis. 2013. MTClassroom and MTDashboard: Supporting Analysis of Teacher Attention in an Orchestrated Multi-Tabletop Classroom. *Int. Conf. on Comput. Supported Collab. Learn.* (2013).

[50] Manolis Mavrikis, Sergio Gutierrez-Santos, and Alex Poulovassilis. 2016. Design and Evaluation of Teacher Assistance Tools for Exploratory Learning Environments. In *Proc. of the 6th Int. Conf. on Learn. Anal. & Knowl.* ACM.

[51] Sharan B Merriam and Elizabeth J Tisdell. 2015. *Qualitative Research: A Guide to Design and Implementation.* John Wiley & Sons.

[52] Inge Molenaar and Carolien Knoop-van Campen. 2019. How Teachers Make Dashboard Information Actionable. *IEEE Trans. on Learn. Technol.* 12, 3 (2019).

[53] Jesús Moreno-León, Gregorio Robles, and Marcos Román-González. 2015. Dr. Scratch: Automatic Analysis of Scratch Projects to Assess and Foster Computational Thinking. *RED. Revista de Educación a Distancia* (2015).

[54] Jeanne Nakamura and Mihaly Csikszentmihalyi. 2014. *The Concept of Flow.* Springer Netherlands.

[55] Daniel G. Oliver, Julianne M. Serovich, and Tina L. Mason. 2005. Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Social Forces* 84, 2 (2005).

[56] John Pane, Elizabeth Steiner, Matthew Baird, and Laura Hamilton. 2015. Continued Progress: Promising Evidence on Personalized Learning. (2015).

[57] Michael Quinn Patton. 2002. *Qualitative Research and Evaluation Methods.* 3rd Sage Publications, Thousand Oaks, CA.

[58] Chris Piech, Jonathan Huang, Andy Nguyen, Mike Phulsuksombati, Mehran Sahami, and Leonidas Guibas. 2015. Learning Program Embeddings to Propagate Feedback on Student Code. In *Proc. of the 32nd Int. Conf. on Mach. Learn.*

[59] Jean Pierce and Deborah Kalkman. 2003. Applying Learner-Centered Principles in Teacher Education. *Theory into Practice* 42, 2 (2003).

[60] Thomas Price, Yihuan Dong, and Dragan Lipovac. 2017. iSnap: Towards Intelligent Tutoring in Novice Programming Environments. In *Proc. of the 2017 ACM SIGCSE Tech. Symp. on Comput. Sci. Educ.* ACM.

[61] Qualtrics. 2022. Fraud Detection. https://www.qualtrics.com/support/survey-platform/survey-module/survey-checker/fraud-detection/

[62] Qualtrics. 2022. Qualtrics XM // The Leading Experience Management Software. https://www.qualtrics.com/

[63] Rebecca Quintana, Chris Quintana, Cheryl Madeira, and James Slotta. 2016. Keeping Watch: Exploring Wearable Technology Designs for K-12 Teachers. In *Proc. of the 2016 CHI Conf. Extended Abstr. on Hum. Factors in Comput. Syst.* ACM.

[64] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, and Yasmin Kafai. 2009. Scratch: Programming for All. *Commun. ACM* 52, 11 (2009).

[65] Gregorio Robles, Jean Carlo Rossa Hauck, Jesús Moreno-León, Marcos Román-González, and Roberto Nombela. 2018. On Tools that Support the Development of Computational Thinking Skills: Some Thoughts and Future Vision. In *Proc. of the Int. Conf. on Comput. Think. Educ.* The Education University of Hong Kong.

[66] María Jesús Rodríguez-Triana, Luis Prieto, Andrii Vozniuk, Mina Shirvani Boroujeni, Beat Schwendimann, Adrian Holzer, and Denis Gillet. 2017. Monitoring, Awareness and Reflection in Blended Technology Enhanced Learning: a Systematic Review. *Int. J. Technol. Enhanc. Learn.* 9, 2-3 (2017).

[67] Olgun Sadik, Anne Ottenbreit-Leftwich, and Thomas Brush. 2020. Secondary Computer Science Teachers' Pedagogical Needs. *Int. J. Comput. Sci. Educ. in Sch.* 4, 1 (2020).

[68] Sue Sentance and Simon Humphreys. 2018. Understanding Professional Learning for Computing Teachers from the Perspective of Situated Learning. *Comput. Sci. Educ.* 28, 4 (2018).

[69] Moin Syed and Sarah C. Nelson. 2015. Guidelines for Establishing Reliability When Coding Narrative Data. *Emerging Adulthood* 3, 6 (2015).

[70] Peeratham Techapalokul. 2017. Sniffing Through Millions of Blocks for Bad Smells. In *Proc. of the 2017 ACM SIGCSE Tech. Symp. on Comput. Sci. Educ.* ACM.

[71] Mike Tissenbaum and James Slotta. 2014. Developing an Orchestrational Framework for Collective Inquiry in Smart Classrooms: SAIL Smart Space (S3). *Int. Conf. the Learn. Sci.* (2014).

[72] Erik van Alphen and Saskia Bakker. 2016. Lernanto: Using an Ambient Display During Differentiated Instruction. In *Proc. of the 2016 CHI Conf. Extended Abstr. on Hum. Factors in Comput. Syst.* ACM.

[73] Elizabeth A Van Es and Miriam Gamoran Sherin. 2002. Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *J. Technol. and Teacher Educ.* 10, 4 (2002).

[74] Anouschka van Leeuwen, Nikol Rummel, Kenneth Holstein, Bruce McLaren, Vincent Aleven, Inge Molenaar, Carolien Knoop-van Campen, Baruch Schwarz, Naomi Prusak, Osama Swidan, et al. 2018. Orchestration Tools for Teachers in the Context of Individual and Collaborative Learning: What Information Do Teachers Need and What Do They Do With It? In *Int. Conf. the Learn. Sci.* Int. Soc. the Learn. Sci.

[75] Jessica Vandenberg, Arif Rachmatullah, Collin Lynch, Kristy E. Boyer, and Eric Wiebe. 2021. Interaction effects of race and gender in elementary CS attitudes: A validation and cross-sectional study. *Int. J. of Child-Comput. Interact.* 29 (2021).

[76] Lev Vygotsky. 1978. Interaction Between Learning and Development. *Read. on the Dev. of Child.* 23, 3 (1978).

[77] David Weintrop. 2019. Block-Based Programming in Computer Science Education. *Commun. the ACM* 62, 8 (2019).

[78] Mark Weiser. 1993. Hot Topics - Ubiquitous Computing. *Computer* 26, 10 (1993).

[79] Mark Weiser and John Seely Brown. 1997. *The Coming Age of Calm Tech.* Springer New York.

[80] Aman Yadav, Sarah Gretter, Susanne Hambrusch, and Phil Sands. 2016. Expanding Computer Science Education in Schools: Understanding Teacher Experiences and Challenges. *Comput. Sci. Educ.* 26, 4 (2016).