CS Principles Goes to Middle School: Learning How to Teach "Big Data"

Philip Sheridan Buffum
Computer Science
North Carolina State Univ.
Raleigh, NC, USA

Allison G. Martinez-Arocho
Computer Science
Meredith College
Raleigh, NC, USA

Megan Hardy Frankosky
Psychology
North Carolina State Univ.
Raleigh, NC, USA

Fernando J. Rodriguez
Computer Science
North Carolina State Univ.
Raleigh, NC, USA

Eric N. Wiebe
STEM Education
North Carolina State Univ.
Raleigh, NC, USA

Kristy Elizabeth Boyer
Computer Science
North Carolina State Univ.
Raleigh, NC, USA

{psbuffum, agmarti5, rmhardy, fjrodri3, wiebe, keboyer}@ncsu.edu

ABSTRACT
Spurred by evidence that students' future studies are highly influenced during middle school, recent efforts have seen a growing emphasis on introducing computer science to middle school learners. This paper reports on the in-progress development of a new middle school curricular module for Big Data, situated as part of a new CS Principles-based middle school curriculum. Big Data is of widespread societal importance and holds increasing implications for the computer science workforce. It also has appeal as a focus for middle school computer science because of its rich interplay with other important computer science principles. This paper examines three key aspects of a Big Data unit for middle school: its alignment with emerging curricular standards; the perspectives of middle school classroom teachers in mathematics, science, and language arts; and student feedback as explored during a middle school pilot study with a small subset of the planned curriculum. The results indicate that a Big Data unit holds great promise as part of a middle school computer science curriculum.

Categories and Subject Descriptors
K.3.2 [**Computers & Education**]: Computer and Information Sciences Education --- *Computer Science Education*

General Terms
Human Factors.

Keywords
Computer science education, middle school, big data

1. INTRODUCTION
Recent years have seen an increased emphasis on introducing computer science to K-12 learners. The motivation for these activities is rooted in an understanding of how students eventually choose and succeed with an undergraduate major: previous positive experience with a subject matter plays a pivotal role [2, 12]. This point has particular significance within the field of computer science; for example, a very small number of US students take the traditional AP CS exam, and there is a troubling participation gap seen in that small body of test-takers [2]. The discrepancy corresponds to similarly disappointing statistics for students who earn bachelor degrees in computer science [29].

In both their stated aspirations and in their development of subject-specific skills, students begin their career trajectory in middle school [14]. Researchers have specifically traced the underproduction and underrepresentation issues in undergraduate computer science departments back to lack of exposure as early as middle school [22, 26]. In the U.S., any middle school computer science intervention must address the challenge that computer science is essentially absent from standard public education curricula. Numerous successful middle school interventions have addressed this challenge through after-school, summer, or enrichment-based activities; for example, these curricula often use various visual programming languages [15] or computer games [25]. Yet important questions remain for a computer science curriculum designed for in-class settings. How can computer science learning objectives be aligned with new curricular standards? How can computer science be successfully delivered by teachers who are trained in other subjects? And finally, how can we build learning activities that are highly engaging and developmentally appropriate for students?

This paper examines these questions in the context of a proposed curricular unit for Big Data, one of the foci of the proposed AP Computer Science Principles course [7]. Although a full formal definition of "Big Data" is still emerging, the phenomenon is often characterized by collection at high *velocity*, storage in a *variety* of structures, and large *volume* [8]. In addition to its widespread societal importance and increasing implications for the computer science workforce, Big Data holds great promise as a focus for middle school computer science because of its rich interplay with other important CS Principles including algorithms, abstraction, and the Internet. As this paper describes, Big Data also represents an area with strong correspondence between CS Principles and emerging US national curricular standards such as the Common Core in Mathematics [5] and the Next Generation Science Standards [17, 18]. In addition to its strong links to other core subjects, the topic of Big Data has great flexibility in

SIGCSE'14, March 5–8, 2014, Atlanta, Georgia, USA.

http://dx.doi.org/10.1145/2538862.2538949

incorporating other highly engaging application domains, including those that are relevant to kids' everyday lives. For instance, an excellent example of Big Data can be seen within social networks and their massive collections of structured and unstructured data.

The work reported here is conducted in the context of the ENGAGE project, a multi-year effort that will see development of a full CS Principles-based middle school curriculum situated within a game-based learning environment. The CS Principles approach may be well suited for the middle school level because the thematic threads of CS Principles Big Ideas can be introduced at a developmentally appropriate level for middle school students and thus provide the basis for more advanced work in high school. We are developing our curriculum in an iterative fashion with pilot studies and focus groups on each set of prototype learning activities. This paper presents the work to date on the Big Data curricular unit, including its alignment with emerging standards, feedback from an interdisciplinary group of teachers, and results from a pilot study with middle school students. While the piloted activity represents only a small subset of the planned curriculum for ENGAGE, results suggest that Big Data holds appeal for achieving computer science learning and for showcasing the relevance of computer science to middle school students.

2. RELATED WORK

Recent years have seen an increase in the number of middle school computing interventions, many of which innovate to teach programming with visual or graphical programming languages. Alice 3D has been used successfully [21], integrating computing within the context of a wide variety of subjects such as math, science, and language arts [20], and to help students understand what their future careers in computing might look like [27]. Scratch programming has also been used extensively [4, 15, 23], including with students with disabilities [1], urban youth [16], or particular underrepresented groups [10]. Other languages such as Logo [15] and more recently Kodu [24] have also been used in middle school interventions. There is also growing momentum for using game design and game programming [25] and robotics [13] as ways to introduce computer science to middle school students. Efforts have also been made to formalize assessment of computational thinking at the middle school level [28].

In complement to this middle school work, the AP Computer Science Principles course is being developed at the high school level and plans to debut as an official, credit-bearing course in the 2016-17 academic year [7]. CS Principles is organized around seven Big Ideas, one of which is Big Data. Students must analyze, interpret, and gain new insight from data sets, both visually and statistically [7]. Relatedly, Berkeley's CS Principles pilot course, "The Beauty and Joy of Computing" has sections on the impact of Big Data and information technology in modern life [11]. Notably, this aspect of social relevance is a motivating factor for women to enter and remain in computer science [9]. Our current work takes inspiration from these CS Principles pilots, in concert with teacher and student focus group feedback, in developing a Big Data curricular unit.

3. BIG DATA CURRICULUM DESIGN

An important focus of our ongoing work is to develop a suite of learning activities that can be adopted within a variety of middle school contexts including standalone computer science courses or as enrichment for subject matter areas including mathematics and science. A vital step toward this goal is aligning the CS Principles-based curriculum with current standards for middle school grade levels. We first analyzed the Common Core State Standards for English and Mathematics, and the Next Generation Science Standards, and then conducted teacher focus groups to gain feedback on this alignment and to help tailor the proposed Big Data curriculum to middle school classrooms.

Table 1. CS Principles Computational Thinking Practices & Common Core Practices for Mathematics

CS Principles Computational Thinking Practices	Common Core Practices for Mathematics
P1: Connecting Computing	MP5: Use appropriate tools strategically
P2: Developing computational artifacts	MP4: Model with Mathematics
P3: Abstracting	MP2: Reason abstractly and quantitatively
P4: Analyzing problems and artifacts	MP1: Make sense of problems and persevere in solving them
P5: Communicating	MP3: Construct viable arguments and critique the reasoning of others MP6: Attend to precision
P6: Collaborating	

3.1 Alignment with Curricular Standards

3.1.1 Common Core State Standards

Beginning in 2009, the Common Core State Standards Initiative created a set of curricular standards intended to align the previously separate curricula of individual states within the U.S. According to its mission statement, these standards, "are designed to be robust and relevant to the real world, reflecting the knowledge and skills that our young people need for success in college and careers" [5]. Common Core for Mathematics laid out a series of high-level practices through which students are expected to find meaning in problems, use abstract and quantitative reasoning, construct arguments, critique the reasoning of others, and model with mathematics. Similarly, the CS Principles curriculum frames seven Big Ideas around the core computational thinking practices of connecting computing, developing computational artifacts, abstracting, analyzing problems and artifacts, communicating, and collaborating [7]. Table 1 highlights some important parallels between the CS Principles Computational Thinking (CT) Practices and the practices for the Common Core in Mathematics.

At a high level, both CS Principles and the Common Core include big ideas on data. While the Common Core standards relating to data do not map directly to the corresponding CS Principles Big Idea, we see the opportunity for a curriculum that serves the goals of both CS Principles and the Common Core in a complementary fashion. Table 2 shows Learning Objectives for CS Principles and selected standards from the Common Core that relate to data.

3.1.2 The Next Generation Science Standards

Similar to the Common Core, the precursor Framework report and Next Generation Science Standards in the U.S. [17, 18] does not explicitly address computer science, but holds promise for the development of complementary curricula. The Framework report states that elements of both computational and mathematical

Table 2. Goals for Data in CS Principles and Common Core

CS Principles Big Idea 3: ***Data and information facilitate the creation of knowledge***
3.1.1 Use computers to process information to gain insight and knowledge
3.1.2 Collaborate when processing information to gain insight and knowledge
3.1.3 Communicate insight and knowledge gained from using computer programs to process information
3.2.1 Use computing to facilitate exploration and the discovery of connections in information
3.2.2 Use large data sets to explore and discover information and knowledge
3.3.1 Analyze the considerations involved in the computational manipulation of information
Common Core Standards that relate to data
Grade 6. SP.A.1 Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers
Grade 6. SP.B.4 Display numerical data in plots on a number line, including dot plots, histograms, and box plots
Grade 6. SP.B.5 Summarize numerical data sets in relation to their context
Grade 7. SP.A.1 Understand that statistics can be used to gain information about a population by examining a sample of the population
Grade 7. SP.A.2 Use data from a random sample to draw inferences about a population with an unknown characteristic of interest.
Grade 7. SP.B.4 Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations
Grade 8. SP.A.1 Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities
Grade 8. SP.A.4 Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table.

thinking are central to K-12 science education, and these are addressed through the discussion of science and engineering practices and crosscutting concepts [18]. These science and engineering practices recognize the importance of identifying student outcomes that go beyond memorizing narrowly focused scientific facts; they provide a broader interpretation of science inquiry that includes computational approaches.

The Next Generation Science Standards stated practices include both analyzing and interpreting data, and using mathematical and computational thinking [17]. The Framework was clear that the eventual Standards documents needed to address how mathematical and computational tools and techniques are deployed in the service of analyzing large data sets: "Such data sets extend the range of students' experiences and help to illuminate this important practice of analyzing and interpreting data" [18]. As this statement illustrates, there is great promise for creating curricula in which the power of computer science is illustrated within the context of widely required subject matter.

3.2 Teacher Focus Groups

With this alignment in hand between the proposed Big Data concepts and standard curricula, the next step was to obtain detailed feedback from experienced middle school teachers to inform the development of specific learning activities. We held a series of in-depth classroom teacher focus groups during the summer of 2013. Three teachers from an urban middle school in Raleigh, North Carolina were recruited to participate for approximately 40 hours each in both group and solo sessions. The teachers were two females and one male, and all three were from minority groups that are underrepresented in computer science. They included one 8th grade math teacher, one 7th grade science teacher, and one 7th grade language arts teacher. Each focus group session lasted between three and four hours and the format varied depending on the topic. Early sessions focused on introducing computer science topics to the teachers. These sessions involved both presentations and guided classroom activities conducted by project team members. Later focus groups were heavily discussion-based, as teachers brainstormed on how to teach computer science principles to middle school students, and drafting classroom activities.

3.2.1 8th Grade Math

The 8th grade math teacher expressed some initial hesitation to the idea of teaching Big Data to middle school students, concerned that its concepts would be out of reach given the students' development level. However, she became much more receptive to the idea after a more thorough discussion and explanation of Big Data. She emphasized the need to introduce topics in "baby steps." She noted that middle school students are still developing the mathematical background to perform statistical computation. A question she raised was, "To what degree and depth do we want them to understand Big Data…and have it support the Common Core?" Another of her primary concerns was, "How do we do it in a way that is not overwhelming to them…and that they say 'Wow, cool. How can I use this in my life today?'" The other two teachers echoed this sentiment, emphasizing that the curriculum should be grounded in human applications to which students can personally relate. As will be described in Section 4, our iterative curriculum development process involves measuring the extent to which students relate to each prototype learning activity.

3.2.2 7th Grade Science

The science teacher provided insight into how to create middle school activities that support what students are learning in their science classes. In particular, she emphasized that we should reinforce the scientific method, which all students in grades six through eight have learned and use in their science classes. She also noted that reading tables and graphs is difficult in general for students, but is an important skill for them to learn. Making inferences or forecasting from data is particularly difficult for middle school students, so she suggested ways to slowly introduce the material over time, and to avoid overly abstract representations. This advice guided the scaffolding included within the first Big Data activity, described in Section 4.

3.2.3 7th Grade Language Arts

The 7th grade language arts teacher quickly saw implications for his classroom. For instance, he uses Wordle[1] to generate "word clouds" from texts before his class reads a book. As a class, they use the most prevalent words to make predictions on the dominant themes that might appear in the book. After the lesson on "What

[1] http://www.wordle.net

is Big Data", the teacher saw this activity as a good opportunity to broaden the discussion in his classroom to include Big Data. Also certified as a social studies teacher, he drew a parallel between a challenge that middle school students might have with programming and one they have in social studies: "sequencing is a difficulty for kids." He noted that middle school students struggle to make the connection from past to present, from present to future, and from self to global community. Our team aims to address some of these challenges with ENGAGE's game-based narrative, designed to give students a strong sense of engagement and personal connection with each problem they solve.

4. PILOT STUDY

Following the conclusion of the focus groups, we conducted a small two-session pilot study with students from an urban middle school in Raleigh, North Carolina. The purpose of the study was to observe students' success at completing a Big Data activity prototype and then to get a sense of their retention and continued interest a week later, as well as have them complete a series of follow-up Big Data learning activities. The first session consisted of a pre-survey, followed by working on the learning activity in collaborative pairs, and then ending with a post-survey. Focus groups held during the second session one week later centered on students' general conceptions surrounding Big Data, as well as their opinions on the various Big Data learning activities.

4.1 Session 1: Big Data Activity

A total of 15 students in 6th and 7th grades participated in the first pilot study session, which was held after school with a total time allotment of 50 minutes including pre-survey and post-survey. Based on experience from prior pilot studies in which students were predominantly male, the after-school coordinator was asked to emphasize recruiting female students. This emphasis succeeded and 13 of the 15 participants were female. The demographic makeup was six white students, four African-American students, one Latino student, one Native American student, one Indian student, one Middle Eastern student, and one multiracial student.

In the pre-survey, students were asked to complete the New General Self Efficacy (NGSE) scale, which aims to explain variance in motivation and performance by measuring students' belief that they can achieve a goal [6]. Scores on the 8-item scale (a 5 point Likert) revealed a high overall level of self-efficacy for the group (M=4.408, SD=0.513). This high overall self-efficacy may reflect a self-selection bias related to the after-school class.

Students also completed Likert scale items scored from 1 (*not at all interested*) to 5 (*very interested*) relating to two application domains: marine life and social networking. Marine life is the application domain for ENGAGE's game-based narrative and an application of Big Data in formal science education. Social networking, on the other hand, is a human application of Big Data that the teacher focus groups indicated would be more relevant to students' everyday lives. On average, students responded to "How interested are you in marine life?" with a 3.1 out of 5 (SD = 1.1), and to "How interested are you in social networks?" with a 3.3 out of 5 (SD = 1.6) While only 2 of the 15 students rated *both* marine life and social networks favorably (4 out of 5, or 5 out of 5), the number of students who rated at least one of these two topics favorably was 12 out of 15, indicating that a Big Data curriculum that includes a variety of topics is much more promising than focusing on just one application domain.

After finishing the online pre-surveys, 37 minutes remained for the Big Data activity prototype. Students were paired at computers

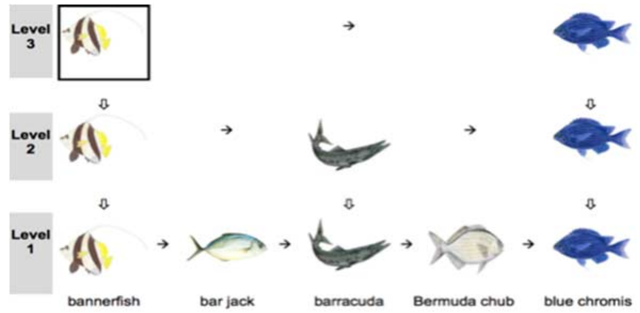

Figure 1. Example visual subset of the fish database

equipped with screen capture and video recording for future analysis. In the activity, students progressed through a written narrative in which they took on the role of computer scientists sent to rescue an underwater research station in distress. As part of the adventure, they had to locate a file for a particular fish species within a massive fish database. The activity guides students step-by-step through a *skip list* search algorithm. Figure 1 shows an example subset of the fish database from the learning activity.

Skip lists, widely used in Big Data applications, provide an excellent context for teaching about data representation and algorithms. While the final curriculum will include many other Big Data learning activities that convey the excitement and conceptual value of Big Data (building upon the success of activities piloted elsewhere [7]), this particular Big Data pilot activity aimed to explore how Big Data applications can provide the framework for addressing other CS Principles Big Ideas, such as computing as a creative activity (Big Idea 1) and algorithms (Big Idea 4). Using a slightly larger subset of the database than the one illustrated in Figure 1, students were tasked with finding the shortest path to various target fish. The activity scaffolded students through each step of the skip list algorithm, using the visual subset of the database. After mastering the searching task manually on that concretely depicted subset, students then used a visual interface (simulating the visual programming language we will use in the ENGAGE game) to arrange pseudo-code blocks into a program that could traverse the entire database to find a desired fish (in this case, Surgeonfish). An ideal algorithmic solution is shown in Figure 2. As described in the next section, this first Big Data activity is designed to occupy one entire classroom session of approximately 50 minutes.

The post-survey included a User Engagement Scale [19] that measured involvement with the activity, focused attention, and the student's desire to repeat a similar activity in the future. Overall, students agreed with the item, "I felt involved with this activity" (3.9 out of 5; SD = 0.9), and generally disagreed with, "I felt discouraged while doing this activity" (1.9 out of 5; SD = 1.0). In contrast to these overall positive ratings, there was one African-

Step 0: Start at **Level 3** Skip List, **Bannerfish**
Step 1: Peek ahead to next fish in current level
Step 2: If the fish you see ahead comes before **<Surgeonfish>**
    Then Move Forward (to that fish)
    And Repeat (return to Step 1)
  Else if the fish you see ahead comes after **<Surgeonfish>**
    Then Move Down (to a lower skip list)
    And Repeat (return to Step 1)
  Else if the fish you see ahead is **<Surgeonfish>**
    Then Move Forward (to that fish)
    And STOP

Figure 2. Target algorithm for students to construct

American female student who responded with a 2 for "I felt involved with the activity", and with a 4 for "I felt discouraged while doing this activity". Among all the students in the study, this student had demonstrated the greatest interest in social networking, with lower stated interest in marine life, possibly explaining her lower feeling of involvement and higher discouragement with the marine life Big Data activity. As will be described in Section 5.2, this student later made contributions during the later focus group that shed light on her experience.

Overall, students rated, "This activity was fun" with 2.8 out of 5 SD = 1.3), lower than hoped. Because the prototype activity is being developed for use within ENGAGE's game-based learning environment, fun is an important emphasis. Upon reflection, we believe that the pace of the learning activity (with several minutes elapsing between achievements on the task) may have hindered a sense of fun, which will be addressed in future iterations. It is important that students frequently feel a sense of achievement, even if the successes represent incremental steps toward the larger goal. Also providing helpful feedback, students did not disagree with "I found this activity to be confusing" (3.2 out of 5; SD = 1.1), indicating room for improvement on the scaffolding. Based on all of the students' feedback, we designed session 2 of the pilot study, which was held one week later.

4.2 Session 2: Focus Groups

Of the original 15 students, 11 returned the next week for the second session, joined by three new students. As with Session 1, all students were in 6th and 7th grades. The demographics for the second session were as follows: 12 female and 2 male; 3 white, 6 African-American, 1 Latino, 1 Native American, 1 Indian, 1 Middle Eastern, and 1 multiracial student. The students were separated into three focus groups of four or five students per group, each facilitated by a member of the research team. Each group spent roughly 15 minutes discussing students' conceptions of computer science and programming, and then 15 minutes on data. A series of follow-up activities to the previous session's Big Data activity were then done in groups, with the new students paired up with students who had attended the previous week. The general purpose of these focus groups was to get student feedback on what works and what needs improvement in our Big Data curriculum. Particular attention was paid to eliciting feedback from students who reported negative opinions on the Big Data activity in session 1.

As predicted, students had a limited understanding of what Big Data might be. They could discuss data on a smaller scale, describing it as information that you can get and store: "anything you want to keep over a long time." They were familiar with examples of data in science, such as weather patterns or plant biology, and some saw the purpose of data being "to do a science project." Although they seemed not to have given much prior thought to other issues related to data, some students got excited thinking about the topic. When asked about data storage, one group of female students interpreted this question in terms of communication and collaboration. Figure 3 shows their conversation.

Facilitator: Where is data stored?
Student 1: On the ground, on a paper, you can put it anywhere.
Student 2: You could put it on the back of a spaghetti can.
Facilitator: Yes you could, but would it be helpful there?
Student 3: You could put [the spaghetti can] in the pantry.
Student 2: Like a community pantry.

Figure 3. Students discuss importance of sharing data

When considered in light of CS Principles Learning Objectives 3.1.2 and 3.1.3 (Table 2), the above dialogue illustrates the potential of data to be socially relevant, a key concern for girls in computer science. Following up on this finding, we also prompted the students to think about data for social networking. Most students had never previously given thought to data outside of science domains, but even students who had been quiet up to this point in the focus group expressed interest in these examples of data in their everyday lives.

After discussing computer science, programming and data, students collaboratively worked on a series of short activities related to the Big Data activity from the session 1. Like that prior activity, these were prototypes of activities being considered for ENGAGE's game-based learning environment. Consequently, they were presented as elements of the game's story, with students told that their feedback would influence the development of the game. In designing these activities, we aimed to improve on the design of the activity from session 1, such as by decreasing the amount of reading required. We also sought to obtain student feedback on Big Data applications that convey the full power and conceptual value of Big Data. Overall, students were noticeably engaged during this session. It included an activity involving cryptography, previously used for engaging middle school students in computer science [3], and a problem in which students solved a mystery by examining clusters of marine data to track a leaked pollutant.

The post-survey revealed that students felt more positively about the activities in this session than they did about the activity from session 1. Students rated these activities highly for "fun" (4.1 out of 5; SD = 0.8) and low for "confusing" (2.1 out of 5; SD = 0.8). Feedback was also more *consistently* positive, including for the African-American female student reporting negative feelings after session 1 (see Section 5.1). During the discussion, she advocated for including social networking features in the ENGAGE game, particularly the ability to chat with other players: "Everyone likes chatting." Similar to the conversation in Figure 3, this student's comments highlight the importance of communication and collaboration, two of the Learning Objectives for CS Principles Big Idea 3 (Table 2). While the improved positivity may be partially explained by repeated exposure to the topic, bringing greater levels of comfort, it also confirms the notion expressed by the classroom teachers in our focus groups: a middle school curriculum can benefit from including activities that are relevant to students' everyday lives.

5. CONCLUSION

A curriculum based on Computer Science Principles holds great promise for middle school, and the CS Principles related to Big Data may be particularly well suited for this age group. We have presented a case for teaching Big Data including its strong alignment with standard U.S. curricula including the Common Core and the Next Generation Science Standards. We have also presented results from classroom teacher focus groups that suggest Big Data can serve to enrich core subject matter including science, mathematics, and language arts. Lastly, we examined the findings of a small pilot study with middle school students that yielded several important lessons for refining the curriculum and learning tasks. First, although we created a learning activity with distinct subtasks, breaking these down even further and providing incremental achievements along the way may increase students' sense of fun and decrease confusion. Second, including a variety of application domains can strategically engage a broader set of learners with the Big Data activities. Finally, focus group feedback indicates that students may intuitively gravitate toward seeing the importance of communication and collaboration, and that these facets may be particularly beneficial to emphasize in Big Data activities for middle school.

To continue creating a successful CS Principles-based Big Data curriculum for middle school, several directions are important to future work. Teacher professional development at the middle school level is key, as computer science and Big Data are new topics for most teachers who might implement the curriculum. Additionally, it is important to develop formal CS Principles-related learning trajectories that map from high school to middle school and eventually to elementary school. Finally, it is hoped that as part of a broader effort to bring CS Principles-based curricula to middle school, this work can substantially increase access to, and interest in, computer science for K-12 learners.

6. ACKNOWLEDGEMENTS

This work is supported in part by the National Science Foundation through Grants CNS-113897 and CNS-1042468. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the participants, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

7. REFERENCES

[1] Adams, J.C. 2010. Scratching Middle Schoolers' Creative Itch. *Proceedings of SIGCSE '10* (2010), 356–360.

[2] Arpaci-dusseau, A. et al. 2013. Computer Science Principles□: Analysis of a Proposed Advanced Placement Course. *Proceedings of SIGCSE '13* (2013), 251-256.

[3] Bergner, N. et al. 2012. Cryptography for Middle School Students in an Extracurricular Learning Place. *CSEDU 2012 Proceedings of the 4th International Conference on Computer Supported Education* (2012), 265–270.

[4] Burke, Q. and Kafai, Y.B. 2012. The Writers ' Workshop for Youth Programmers Digital Storytelling with Scratch in Middle School Classrooms. *Proceedings of SIGCSE '12* (2012), 433–438.

[5] CCSSI, Common Core State Standards Initiative: 2012. Retrieved September 05, 2013 from *http://www.corestandards.org/Math*

[6] Chen, G. et al. 2001. Validation of a New General Self-Efficacy Scale. *Organizational Research Methods*. 4, 1 (Jan. 2001), 62–83.

[7] Computer Science: Principles: 2013. Retrieved September 5, 2013 from *http://www.csprinciples.org/*

[8] Crawford, K. 2011. Six Provocations for Big Data. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society* (2011), 1–17.

[9] Fisher, A. et al. 1997. Undergraduate Women in Computer Science: Experience, Motivation and Culture. *Computer*. 29, (1997), 106–110.

[10] Franklin, D. et al. 2011. Animal Tlatoque: Attracting Middle School Students to Computing Through Culturally-Relevant Themes. *Proceedings of SIGCSE '11* (2011), 453–458.

[11] Garcia, D.D. et al. 2012. Rediscovering the Passion, Beauty, Joy, and Awe: Making Computing Fun Again, part 5. *Proceedings of SIGCSE '12* (2012), 577–578.

[12] Guzdial, M. et al. 2010. A Statewide Survey on Computing Education Pathways and Influences□: Factors in Broadening Participation in Computing. *Proceedings of the ninth annual international conference on International computing research (ICER '12)* (2012), 143–150.

[13] Larkins, D.B. et al. 2013. Application of the Cognitive Apprenticeship Framework to a Middle School Robotics Camp. *Proceedings of SIGCSE '13* (2013), 89–94.

[14] Lent, R.W. et al. 1994. Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *Journal of Vocational Behavior*. 45, (1994), 79–122.

[15] Lewis, C.M. 2010. How Programming Environment Shapes Perception, Learning and Goals: Logo vs. Scratch. *Proceedings of SIGCSE '10* (2010), 346–350.

[16] Maloney, J. et al. 2008. Programming by Choice□: Urban Youth Learning Programming with Scratch. *Proceedings of SIGCSE '08* (2008), 367–371.

[17] Next Generation Science Standards: 2013. Retrieved September 5, 2013 from *http://www.nextgenscience.org/next-generation-science-standards*

[18] NRC, National Research Council. 2011. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The National Academies.

[19] O'Brien, H.L.O. and Toms, E.G. 2010. The Development and Evaluation of a Survey to Measure User Engagement. *Journal of the American Society for Information Science*. 61, (2010), 50–69.

[20] Rodger, S.H. et al. 2009. Engaging Middle School Teachers and Students with Alice in a Diverse Set of Subjects. *Proceedings of SIGCSE '09* (2009), 271–275.

[21] Rodger, S.H. et al. 2010. Enhancing K-12 Education with Alice Programming Adventures. *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education - ITiCSE '10*. (2010), 234.

[22] Shashaani, L. 1994. Gender Differences in Computer Experiences and its Influence on Computer Attitudes. *Journal of Educational Computing Research*. 11, (1994), 347–367.

[23] Sivilotti, P.A.G. and Laugel, S.A. 2008. Scratching the Surface of Advanced Topics in Software Engineering: A Workshop Module for Middle School Students. *Proceedings of SIGCSE '08* (2008), 291–295.

[24] Stolee, K.T. and Fristoe, T. 2011. Expressing computer science concepts through Kodu game lab. *Proceedings of SIGCSE '11* (2011), 99–104.

[25] Webb, D.C. et al. 2012. Toward an Emergent Theory of Broadening Participation in Computer Science Education. *Proceedings of SIGCSE '12* (2012), 173–178.

[26] Webb, H.C. 2011. Injecting Computational Thinking into Career Explorations for Middle School Girls. *2011 IEEE Symposium on Visual Languages and HumanCentric Computing VLHCC*. (2011), 237–238.

[27] Webb, H.C. and Rosson, M.B. 2011. Exploring Careers While Learning Alice 3D: a Summer Camp for Middle School Girls. *Proceedings of SIGCSE '11* (2011), 377–382.

[28] Werner, L. et al. 2012. The Fairy Performance Assessment: Measuring Computational Thinking in Middle School. *Proceedings of SIGCSE '12* (2012), 215-220.

[29] Zweben, B.S. 2013. Computing Degree and Enrollment Trends From the 2011-2012 CRA Taulbee Survey. Computing Research Association.