# An Analysis through an Equity Lens of the Implementation of Computer Science in K-8 Classrooms in a Large, Urban School District

Jean Salac, Max White, Ashley Wang, Diana Franklin
University of Chicago
Chicago, Illinois
{salac,mnwhite,awang21,dmfranklin}@uchicago.edu

## ABSTRACT

Major metropolitan school districts around the United States are implementing computer science in elementary school classrooms as part of the CS for All (CS4All) initiative. Little is known, however, about the success of such a large-scale rollout, especially in terms of equity. In this study we analyze the performance of 4th grade classrooms completing three modules of an introductory computational thinking curriculum, looking at not only overall results but also the variance in performance between high-, mid-, and low-performing schools (as identified by their school report cards). We find that all classrooms are benefiting from the computational thinking (CT) curriculum, making great strides in providing equitable access to CT education. However, statistically-significant differences in performance are present, especially between the high- and low-performing schools, showing that there is still room for improvement in developing strategies and curricula for struggling learners.

## CCS CONCEPTS

• **Social and professional topics → Computer science education**; *Computational thinking*; *K-12 education*;

## KEYWORDS

K-8 education, diversity, Scratch

**ACM Reference Format:**
Jean Salac, Max White, Ashley Wang, Diana Franklin. 2019. An Analysis through an Equity Lens of the Implementation of Computer Science in K-8 Classrooms in a Large, Urban School District. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19), February 27-March 2, 2019, Minneapolis, MN, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3287324.3287353

## 1 INTRODUCTION

The CS4All movement in the United States has gained momentum in the past decade. Its genesis might be the landmark research and interventions in the Los Angeles Unified School District chronicled in Stuck in the Shallow End [14] coupled with the National Science Foundation CS10K program's goal of training ten thousand teachers to teach computer science in high schools across the country [1].

Several public school districts, including Chicago, New York City, and San Francisco, have added computer science (CS) graduation requirements, with accompanying teacher professional development (PD). This provides fertile ground to understand the impacts of CS efforts on student learning.

This study analyzes the equity of CS learning outcomes across three different-performing schools in the district. School performance has been shown to be a proxy for the race, income, and parental involvement of their students [13, 15] and related to resources and teacher turnover rates [2]. Such disparities are realities in school districts nationwide, and thus need to be addressed to advance equity in K-12 CS education.

This study focuses on a large urban school district's nascent CS4All effort, which will provide CS instruction to all pre-K-12th grade students. This study focuses on 4th grade, using an adaptation of the Creative Computing Curriculum [5] that teaches sequence, events, and iteration.

Our study was guided by the following research questions:

- What is the level of understanding and sources of confusion overall in learning the core concepts of events and iteration?
- To what extent does school performance influence results?

This paper presents key findings from assessments given at the conclusion of each module. In the next section, we present relevant literature on large CS4All initiatives and student learning. In Section 3, we describe our methods and experimental design. We present our results in Section 4, along with a discussion of the implications of the results. Section 5 describes limitations of this study. Finally, we provide concluding remarks in Section 6.

## 2 RELATED WORK

We present two bodies of work that overlap with this study's goals. First, we present studies that have added to the collective understanding of large-scale efforts to improve computer science education. Second, we present studies that focus on student learning of basic computing concepts.

Margolis et al. began research on interventions to improve K-12 equity in LAUSD [14]. They focused on structural barriers, curriculum shortcomings and professional development shortcomings. This heavily influenced CS10K, a movement to develop high school curricula and train ten thousand teachers. A national landscape study was performed for CS10K to understand the professional development opportunities for teachers in the United States [8],


*SIGCSE '19, February 27-March 2, 2019, Minneapolis, MN, USA*

ACM ISBN 978-1-4503-5890-3/19/02…$15.00
https://doi.org/10.1145/3287324.3287353

identifying gaps in offerings at the time. Wang et al. surveyed households and schools, finding that while there was a high demand for CS education from both parents and students, over 75% of principals reported their school did not offer computer science with programming/coding [17].

Major school districts have published early results of CS4All efforts. A study in New York City found that schools with CS courses and activities served fewer Black and Latino students and more White and Asian students, compared with schools without CS courses [6]. Research in Broward County in Florida found an association between completing more Code.org computer science lessons and higher literacy scores [3].

A growing body of work has researched challenges and successes students have had using visual block-based languages (VBBLs) [12] as they learn initialization [9], variables and loops [11], what concepts are appropriate for what age groups [7, 10, 16], and the patterns that emerge from starting with Scratch that might impact later instruction [18].

This work reinforces prior CS research in novice learning in two ways. First, it provides insight into equity at a district level with respect to student learning. Second, it is a larger-scale study researching student learning in the concepts of sequence, events, and repetition.

## 3 METHODS

This study consisted of nine 4th grade classrooms (204 9-10 year old students): three classrooms from three schools (identified as high-, mid-, and low-performing by the school district) in a large urban school district. In their classification, they use characteristics of both students (e.g. percentage of minority students, English language learners, students with special needs, students in poverty, etc) and teachers (e.g. years of experiences, turnover rates, etc).

All teachers in the study underwent the same professional development and taught the same curriculum. In each school, the same teacher taught all three classrooms. Each school has balanced gender splits. Ethnic breakdowns are shown in Table 1, matching prior research that underrepresented minorities are over-represented in mid- and low-performing schools [13, 15], allowing equity in terms of overall school performance to function as a proxy for equity in terms of minority status.

| School | Ethnicity (%) | | | | |
|---|---|---|---|---|---|
| | Asian | Black | His/Lat | Pac Is | White |
| High | 49 | 0 | 13 | 2 | 21 |
| Mid | 2 | 14 | 54 | 5 | 18 |
| Low | 5 | 8 | 26 | 31 | 0 |

**Table 1: Demographics of study schools.**

Students completed three modules in an introductory CT curriculum in Scratch, which was a modification of the Creative Computing Curriculum [5]. Upon completion of Unit 2 (events) and Unit 3 (loops), students took a pen-and-paper assessment designed for this curriculum by a team of researchers, practitioners, and a reading comprehension specialist, with several rounds of review by each of the five team members. Each assessment consisted of a mix of multiple-choice, fill-in-the-blank and open response questions, and was designed to take students about 20 to 30 minutes to complete. Multiple-choice answer options were scrambled randomly on each exam. The assessments were graded by undergraduate researchers. The open response question was qualitatively coded by two undergraduate researchers to ensure inter-rater reliability.

Our study follows the completely randomized hierarchical *CRH-pq(A)* design. The linear model is as follows:

$$Y_{ijk} = \mu + \alpha_j + \beta_{k(j)} + \epsilon_{i(jk)} \quad (1)$$

where:

- $Y_{ijk}$ is the question score for the $i^{th}$ student in classroom $k$ within school $j$,
- $\mu$ is the grand mean of the question score,
- $\alpha_j$ is the effect of school $j$,
- $\beta_{k(j)}$ is the effect of classroom $k$ within school $j$,
- and $\epsilon_{i(jk)}$ is the error effect associated with $Y_{ijk}$.

The independent variable in this study is the type of school, with classrooms nested within them. Both the type of school and individual classrooms are fixed factors. The classrooms in our study are of different sizes, so we randomly sampled classrooms of 18 students (the smallest classroom size in our study) and ran the linear model based on the sampled classrooms. This process was repeated 1000 times, and the average of the linear model outputs over all iterations was calculated; the mean of the outputs was used.

Because there are three schools, analysis was performed in two steps to find statistical significance. First, an ANOVA F-test was used to find whether there are any statistically-significant differences between schools. Then, a Fisher-Hayter Post Hoc test was performed pairwise on the three pair choices to determine which pairs' result differences were statistically significant. Both tests provide $p$ values — $p < 0.05$ is statistically significant. The ANOVA F-test also provides an $F$ value for a significance test, which is compared to values in the $F$ distribution. Effect sizes ($\omega^2$) were also calculated, where $\omega^2$ values of 0.010, 0.059, and 0.138 are small, medium, and large associations, respectively [4]. With the large sample size, the power of all tests was at least 80%.

## 4 RESULTS

Analysis seeks to answer two questions, one about overall learning and confusion with events (Q1-2) and loops (Q3-7,EC) and one about the correlation between school performance and student learning. Question-level results are presented, both overall and per-school, along with discussion of what those results reveal about understanding within the concept. We then discuss overall implications of the study in the discussion.

### 4.1 Q1: Events Starting One Script

Question 1 asked students to circle which script(s) out of the four shown would run if they clicked on the sprite. Two scripts started with 'when sprite clicked', one with 'when green flag clicked', and one with 'when space key pressed'. Students received two points for every correct script circled and lost one for any incorrect script circled, for 0-4 points.

The overall average score on Q1 was 2.49 (Figure 1). 61.4% of students circled only correct scripts, but only 43.4% of students

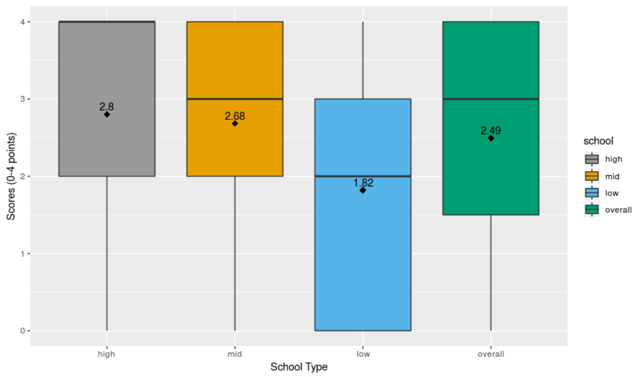


**Figure 1: Q1 Events Scores(0-4 points)**

circled all correct scripts (and no incorrect ones). Across all three schools, there is a statistically-significant difference ($F(2,144)=7.43$, $p < 0.001$, $\omega^2=0.282$). Between pairs of schools, there are significant differences between the low-performing school and both the high- and mid-performing schools with a Fisher-Hayter Post Hoc ($p < 0.05$).

To better understand how students answered, student responses are categorized as: (1) NO correct - students who circled none of the correct answers, (2) BOTH correct & incorrect - students who circled some some correct and some incorrect answers, (3) ONLY correct - students who circled correct but not incorrect answers, and (4) ALL correct & NO incorrect - students who circled all the correct answers and none of the incorrect ones. As shown in Table 2, students in the high-performing school circled correct options (ONLY C) most frequently and provided the most complete answers, followed by the mid- and low-performing schools. Conversely, students in the low-performing school circled incorrect options (NO C, Both C/I) most frequently and were most likely to miss correct options, followed by the mid- and high-performing school.

| Sch | Category | | | |
|---|---|---|---|---|
| | NO C | Both C/I | ONLY C | ALL C/NO I |
| H | 19.3% | 7.9% | 72.7% | 59.1% |
| M | 24.5% | 15.9% | 59.5% | 41.1% |
| L | 33.3% | 16.7% | 50% | 26.7% |
| All | 24.8 | 13.8% | 61.4% | 43.4% |

**Table 2: Q1 Qualitative Results**

## 4.2 Q2: Events Starting Multiple Scripts

Question 2 consists of two actions (playing drum and changing costume) in three scripts across two sprites (Pico & Giga), all started by `when green flag clicked`. Pico's single script performs the actions sequentially, whereas Giga's two scripts run in parallel (Figure 2). To assess students' understanding of multiple events in multiple scripts versus sequential events in one script, students were asked to circle the true statements from the following:

a) Pico plays the drum 7 times THEN changes costumes 4 times.
b) Giga plays the drum 7 times THEN changes costumes 4 times.

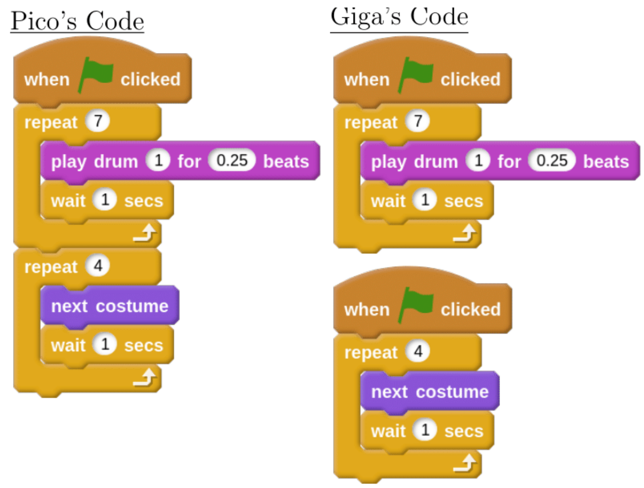


**Figure 2: Q2 Sequential (left) and Parallel (right) Scripts**

c) Pico plays the drum AND changes costumes at the same time.
d) Giga plays the drum AND changes costumes at the same time.
e) Pico and Giga both play the drum 7 times THEN change costumes 4 times.

The correct answers were a and d. Students earned 2 points for each correct answer circled and lost 1 point for each incorrect answer circled, for 0-4 points. Most students struggled with Q2, with an overall average score of 1.11 points. When broken down by school, the average scores were 1.31, 1.4 and 0.53 points for high-, mid-, and low-performing schools, respectively. Across all three schools, there is a statistically-significant difference ($F(2,144)=7.82$, $p < 0.001$, $\omega^2=0.289$). Between pairs of schools, there are significant differences between the low-performing school and both the high- and mid-performing schools with a Fisher-Hayter Post Hoc ($p < 0.05$).

64.37%, 70.13%, and 46.55 % of students in high-, mid-, and low-performing schools, respectively, correctly identified Pico's sequential behavior. Only 41.38%, 36.91%, and 35.79% of students in high-, mid-, and low-performing schools, respectively, however, circled Giga's parallel behavior.

Some very common errors include: 44.82% circled Giga having sequential behavior; 22.07% circled Pico having parallel behavior; and 53.85% circled option 5 (both sprites have sequential behavior). The higher frequency of answers with sequential behavior suggest that students may not understand parallelism as deeply as sequential execution in Scratch.

However, students may also have had difficulty comprehending the answer choices – 25.75% of students selected incompatible answers (such as circling both a and c or b and d). Determining whether the difference in performance was due to different levels of CS understanding or different reading abilities is further study.

## 4.3 Q3: Repeat Iteration Count

Students were shown a repeat block and asked how many times the loop would repeat.

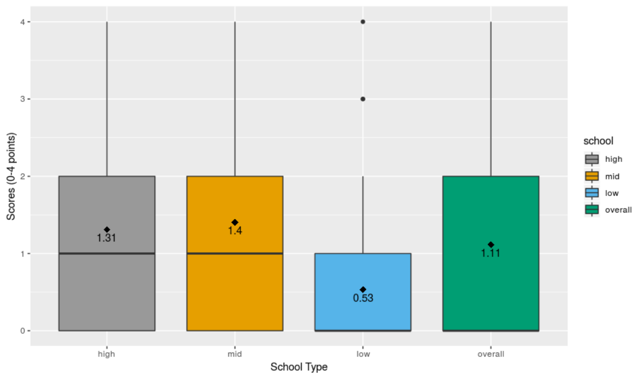


**Figure 3: Q2 Parallel Scripts Scores (0-4 points)**

Almost all of the students from each school were able to answer correctly, with 98.85%, 88.31%, and 84.48% of the students in the high-, mid-, and low-performing schools, respectively getting the answer correct (Figure 4). Comparing the differences in the number of students who answered correctly, we found a statistically-significant difference (F(2,144)=5.05, p <0.01, $\omega^2$=0.192) among the schools. A Fisher-Hayter Post Hoc pairwise analysis shows a significant difference between the high- and low-performing schools (p < 0.05).

Students performed very well on Q3. However, there is still a performance gap between the high- and low-performing schools.

## 4.4 Q4: Unrolling a Loop

Students were shown a repeat 4 loop consisting of two blocks.They were given choices of those two blocks repeated 1, 2, 3, and 4 times. Students were then asked to choose the unrolled code that did the same thing as the loop.

Students struggled on Q4, with only 56.44% overall answering it correctly. Within individual schools, 70.11%, 53.05%, and 44.83% of the students in the high-, mid-, and low-performing schools, respectively, answered correctly (Figure 4). There is a statistically-significant difference among schools for Q2 (F(2,144)=5.25, p <0.01, $\omega^2$=0.214), with only a significant difference between high- and low-performing schools from a Fisher-Hayter Post Hoc (p <0.05).

When we put Q3 and Q4 performance in perspective, we see that while students are able to identify how many times a repeat loop is run, many students do not truly understand what that means. This implies a limited understanding of loop functionality, especially in the low-performing school.

## 4.5 Q5: Repeated Blocks vs Repeat Loops

Students were asked to circle the scripts that would make a sprite perform some actions exactly three times. Students were provided one set of blocks (a) alone and (b) inside a repeat 3 loop, and three sets of sequential blocks (c) alone and (d) within a repeat block (Figure 5). Q5 was designed based on a common misconception observed by teachers - not understanding the relationship between repeated code within a loop and repeated loop iterations. Choices were provided in random order on different assessments.

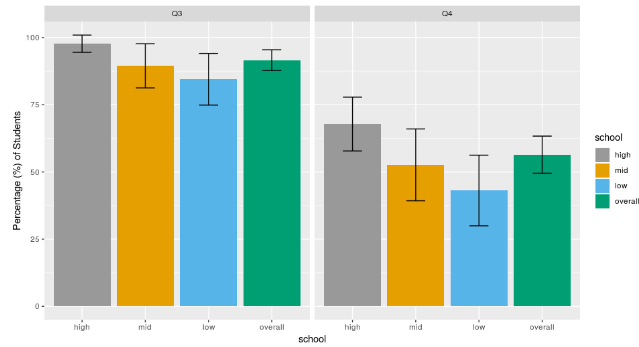


**Figure 4: Q3 Iteration Count vs Q4 Loop Unrolling Results**


**Figure 5: Q5 Answer Option (d) and inspiration for question.**

Q5 had two correct answers (b and c described above); students received two points for each correct answer circled and lost one point for each incorrect answer circled, for 0-4 points. Overall, the median score was 2 and mean score was 2.07. Comparing performance over all three schools, we found a significant difference (F(2,144)=7.00, p < 0.01, $\omega^2$=0.272). A Fisher-Hayter Post Hoc revealed a significant difference between the high-performing school and the other two schools (p < 0.05).

Of the incorrect options, the option with only one pair of blocks was the least common, with 3.45%, 5.15%, and 1.72% of students choosing that option in the high-, mid-, and low-performing schools, respectively. In contrast, 25.29%, 28.57%, and 44.83% of students in high-, mid-, and low-performing schools selected the option with three pairs of blocks within a 'Repeat 3' block, supporting the observation made by the teachers.

## 4.6 Q6: Loops Within Sequence

Question 6 consisted of a repeat loop sandwiched between two blocks and asked them three sub-questions: which blocks run (a) *in*, (b) *before*, and (c) *after* the loop. On each sub-question, students earned 2 points for each correct answer circled and lost 1 point for each incorrect answer circle, for 0-4 points (a) or 0-2 points (b, c).

We found a significant difference in scores across schools for all three parts (a: F(2,144)=7.00, p < 0.01, $\omega^2$=0.289; b: F(2,144)=7.78, p <0.001, $\omega^2$=0.292; c: F(2,144)=9.40, p < 0.001, $\omega^2$=0.329). A Fisher-Hayter Post Hoc shows a significant difference between the high-performing school and the other two schools (p < 0.05).

Students from the high-performing school outperformed the rest of the students on all parts of Q6. Part (a) tests their understanding of

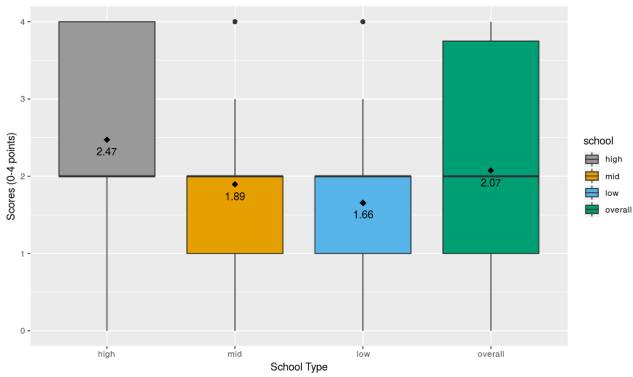


**Figure 6: Q5: Repeated Blocks vs Loops Scores (0-4 points)**

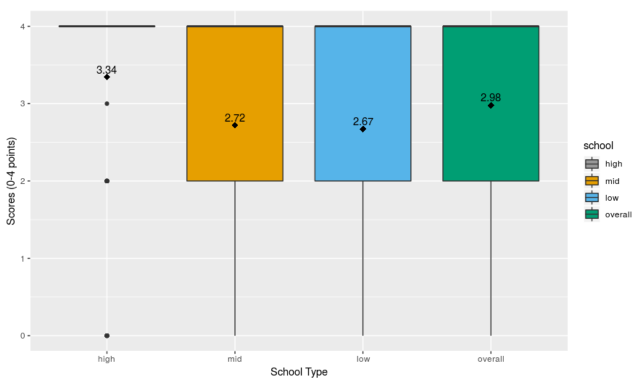


**Figure 7: Q4(a) Code in Loop Scores (0-4 points)**

loops, while parts (b) and (c) test their understanding of sequence, a topic that was covered earlier in the curriculum. This result suggests the students from the high-performing school were better able to retain CT concepts.

However, two similar-looking blocks used in Q6 may have unintentionally confused students in part c. We found that the `play drum` block, which looks similar to the correct `play sound` block, was a frequent incorrect answer, selected by 9.2%, 11.04% and 27.59% of students in the high-, mid-, and low-performing school, respectively. Future work will explore whether reading comprehension skills may have influenced their performance on Q6.

## 4.7 Q7: Explain In Your Own Words

Students were shown a loop and asked to explain what the loop would do in their own words. Answers were given between 0-10 points depending on accuracy and completeness.

Students performed fairly well, with an average score of 8.36. When broken down by school, the average scores were 9.33, 8.04 and 7.21 points for high-, mid-, and low-performing schools, respectively. Across all three schools, there is a statistically-significant difference ($F(2,144)=24.63$, $p < 10^{-11}$, $\omega^2=0.524$). There are statistically-significant differences between all pairs of schools from a Fisher-Hayter Post Hoc ($p < 0.05$).

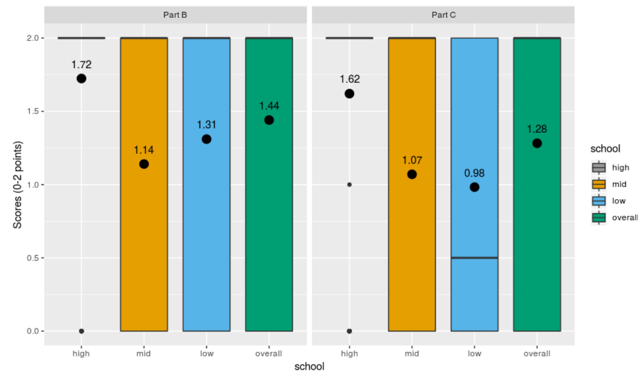


**Figure 8: Q4 Code before (b) and code after (c) Loop Scores (0-2 points)**

Qualitative coding of student responses reveals some patterns. In the high-performing school, all students described the loop itself and at least one of the blocks in the loop. In the mid-performing school, 5.19% described neither the loop nor any blocks within it, and 1.30% did not describe any blocks in the loop. In the low-performing school, these figures jump to 10.1% for the former feature and 6.74% for the latter feature.

While both features were less frequent in high- and mid-performing schools, other answer features were more prevalent in their responses. 14.04% and 16.3% of students in the high- and mid-performing schools, respectively, left out block parameters in their responses, compared with 10.11% of students in the low-performing school. 14.04% and 20% of students in the high- and mid-performing schools, respectively, neglected to describe at least one block in their responses, compared with 6.74% of students in the low-performing school. 17.5% and 9.09% of students in the high- and mid-performing schools, respectively copied the text from the Scratch blocks verbatim, compared with 3.37% of students in the low-performing school.

These patterns in answer features suggest that, in the low-performing school, struggling students answered minimally. In contrast, struggling students in the high- and mid-performing schools attempted to write correct answers, but would either leave parts out or copy them directly from the Scratch code. Determining whether these patterns are due to different levels of CS understanding, or different reading and writing abilities will require further study.

## 4.8 EC: Nested Loop Iteration Count

The last problem, an Extra Challenge question, presented a nested loop, which was not explicitly taught in the curriculum. It consisted of a repeat 2 outer loop and a `repeat 10` inner loop, and we asked students how many times the blocks in the inner loop would run.

Overall, only 25.25% of students gave the correct answer of 20. When broken down by school, we found that 38.76%, 24.08%, and 8.62% of students answered correctly in the high-, mid-, and low-performing schools, respectively. Comparing across schools, there is a statistically-significant difference in performance ($F(2,144)=7.85$, $p < 0.001$, $\omega^2=0.285$). Between pairs of schools, students in both the

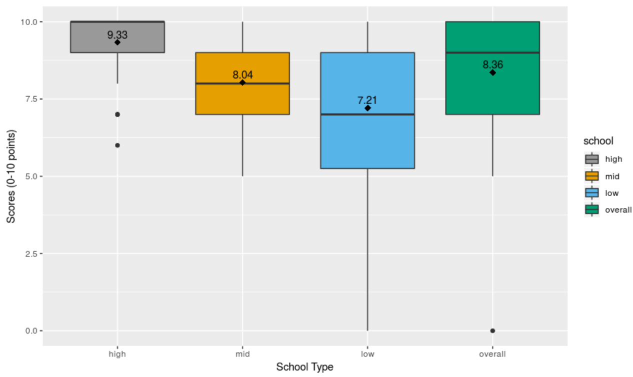


Figure 9: Q7 Explain In Your Own Words Scores(0-10 points)

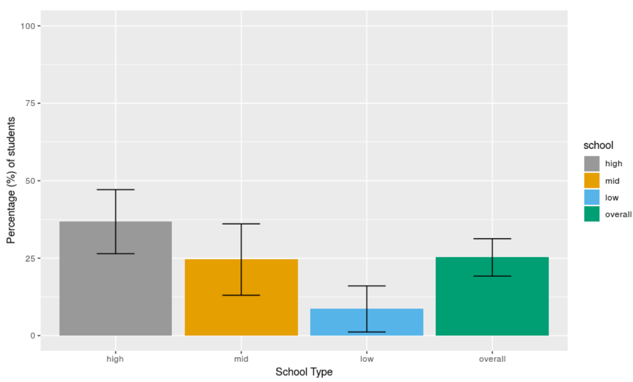


Figure 10: EC Nested Loop Results

high- and mid-performing schools significantly outperformed students from the low-performing school from a Fisher-Hayter Post Hoc ($p < 0.05$).

Our results suggest that students in the high- and mid-performing schools were more likely to apply what they learned about repetition to nested loops, compared to students in the low-performing school. Nonetheless, two responses displayed an understanding of basic loops, if not nested loops. A plurality (40.13%) of students wrote 10, the number in the innermost loop, and 7.36% of students wrote 2, the number in the outermost loops.

### 4.9 Discussion

We now revisit our two original research questions:

- What is the level of understanding and sources of confusion overall in learning the core concepts of events and iteration?
- To what extent does school performance influence results?

Students in high-performing schools, in general, show a good understanding of events and loops. Most students know the number of iterations a repeat loop performs (Q3), can see the relationship between the loop and equivalent sequential code (Q4, Q5), understand the order of blocks in a loop compared to blocks before and after the loop (Q6), and can articulately describe what a loop does with fidelity (Q7). Only two concepts, parallelism (Q2) and nested loops (EC) were beyond their grasp.

| Q | Mean | | | Significant Differences | | |
|---|---|---|---|---|---|---|
| | High | Mid | Low | H*M | H*L | M*L |
| Q1 (4) | 2.80 | 2.69 | 1.84 | | * | * |
| Q2 (4) | 1.31 | 1.4 | 0.53 | | * | * |
| Q3 (1) | 0.99 | 0.88 | 0.84 | | * | |
| Q4 (1) | 0.70 | 0.53 | 0.45 | | * | |
| Q5 (4) | 2.47 | 1.89 | 1.56 | * | * | |
| Q6a (4) | 3.34 | 2.72 | 2.67 | * | * | |
| Q6b (2) | 1.72 | 1.14 | 1.31 | * | * | |
| Q6c (2) | 1.62 | 1.07 | 0.98 | * | * | |
| Q7 (10) | 9.33 | 8.04 | 7.21 | * | * | * |
| EC (1) | 0.39 | 0.24 | 0.09 | | * | * |

Table 3: Summary of student performance and statistically-significant differences between schools.

However, our results show that students at mid- and low-performance schools exhibit a much shallower understanding of loops. While they can specify how many times a repeat loop will iterate (86%), fewer than half can identify the unrolled equivalent of a repeat loop and identify both constructs that repeat actions (repeat loop and sequential code). Overall, there was a statistically-significant difference between students in the high- and low-performing school on all questions, and the mid- and low-performing schools on Q5, Q6, and Q7.

In terms of ethnic equity, the high-performing school contains high percentages of students already well-represented in the computing field (Asian and White), whereas the majority of Black, Hispanic/Latino, and Pacific Islander students were at mid- and low-performing schools.

## 5 LIMITATIONS

There are several limitations to this study. First, due to the lack of validated assessments for this age group, this assessment was created specifically for this curriculum, and is therefore not validated. However, our design team included practitioners, CS Ed researchers, and a specialist in students with disabilities. Also, the classrooms, students, and teachers were not randomly sampled throughout the individual school district, nor school districts in the US or world. The same teacher taught all three classrooms within each school; therefore differences could be attributed to teaching methods. Finally, the demographics of study schools may not hold for other examples of high-, mid-, and low-performing schools. More research on a large sample set is necessary to understand district- and nation-wide implications of such results.

## 6 CONCLUSION

This study shows the successes and challenges of providing elementary school computational thinking instruction with the goal of providing equitable outcomes. This curriculum supports some students in learning sequence, events, and loops well. However, special care should be taken to improve the depth of understanding in mid- and low-performing schools, especially for loops.

## REFERENCES


[1] Owen Astrachan et al. "The CS10K project: mobilizing the community to transform high school computing". In: *Proceedings of the 42nd ACM technical symposium on Computer science education*. ACM. 2011, pp. 85–86.

[2] Jack Buckley, Mark Schneider, and Yi Shang. "Fix it and they might stay: School facility quality and teacher retention in Washington, DC". In: *Teachers College Record* 107.5 (2005), pp. 1107–1123.

[3] Jeanne Century, Kaitlyn Ferris, and Huifang Zuo. *Finding Time for Computer Science in the Elementary Day*. URL: http://outlier.uchicago.edu/TimeforCS/findings/.

[4] Jacob Cohen. *Statistical power analysis for the behavioural sciences*. 1988.

[5] Creative Computing. *An introductory computing curriculum using Scratch*.

[6] Cheri Fancsali et al. "A Landscape Study of Computer Science Education in NYC: Early Findings and Implications for Policy and Practice". In: *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. ACM. 2018, pp. 44–49.

[7] Louise P Flannery et al. "Designing ScratchJr: support for early childhood learning through computer programming". In: *Proceedings of the 12th International Conference on Interaction Design and Children*. ACM. 2013, pp. 1–10.

[8] Baker Franke et al. "Expanding access to K-12 computer science education: research on the landscape of computer science professional development". In: *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM. 2013, pp. 541–542.

[9] Diana Franklin et al. "Initialization in scratch: Seeking knowledge transfer". In: *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM. 2016, pp. 217–222.

[10] Diana Franklin et al. "Using upper-elementary student performance to understand conceptual sequencing in a blocks-based curriculum". In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. ACM. 2017, pp. 231–236.

[11] Shuchi Grover and Satabdi Basu. "Measuring student learning in introductory block-based programming: Examining misconceptions of loops, variables, and boolean logic". In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. ACM. 2017, pp. 267–272.

[12] Charlotte Hill et al. "Floors and Flexibility: Designing a programming environment for 4th-6th grade classrooms". In: *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. ACM. 2015, pp. 546–551.

[13] Jennifer Jellison Holme. "Buying homes, buying schools: School choice and the social construction of school quality". In: *Harvard Educational Review* 72.2 (2002), pp. 177–206.

[14] Jane Margolis. *Stuck in the shallow end: Education, race, and computing*. MIT Press, 2010.

[15] Sean F Reardon. "The widening income achievement gap". In: *Educational Leadership* 70.8 (2013), pp. 10–16.

[16] Linda Seiter and Brendan Foreman. "Modeling the learning progressions of computational thinking of primary grade students". In: *Proceedings of the ninth annual international ACM conference on International computing education research*. ACM. 2013, pp. 59–66.

[17] Jennifer Wang et al. "Landscape of K-12 computer science education in the US: Perceptions, access, and barriers". In: *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM. 2016, pp. 645–650.

[18] David Weintrop et al. "Starting from Scratch: Outcomes of Early Computer Science Learning Experiences and Implications for What Comes Next". In: *Proceedings of the 2018 ACM Conference on International Computing Education Research*. ACM. 2018, pp. 142–150.