

Article

Where Socioeconomic Differences in Computational Thinking Become Visible: Integrating Diagnostic and Log-Based Behavioral Assessment

Ben Avital-Lev *  and Arnon HersHKovitz 

School of Education, Tel Aviv University, Tel Aviv 6997801, Israel; arnonhe@tauex.tau.ac.il

* Correspondence: benavital@mail.tau.ac.il

Abstract

This study examines where socioeconomic differences in students' computational thinking (CT) learning become visible by comparing a diagnostic assessment of conceptual CT knowledge with behavioral indicators derived from interaction data in a digital programming environment. The study involved 444 elementary school students who completed a structured sequence of programming tasks while their activity was recorded. Conceptual CT knowledge was assessed using a validated diagnostic instrument, and four behavioral indicators were derived from learning logs: average first-try stars, attempts per challenge, highest challenge reached, and average solution time. Analyses were conducted at two complementary levels: individual indicators and integrated digital behavioral types identified through clustering. The findings revealed no meaningful socioeconomic differences in diagnostic CT performance and no consistent differences across most individual behavioral indicators, with the exception of average first-try stars. However, socioeconomic differences became visible when students' interaction patterns were examined as multidimensional configurations of engagement. These results suggest that socioeconomic variation is reflected primarily in students' engagement with digital problem-solving processes rather than in conceptual knowledge alone. The study highlights the value of combining diagnostic and log-based measures for understanding how educational inequality may become observable in computational thinking development.

Keywords: computational thinking; socioeconomic status; educational inequality; learning analytics; programming education; log-based assessment; digital learning



Academic Editors: Costas Constantinou, Lisa Dikomitis and Eirini Kampriani

Received: 25 March 2026

Revised: 4 June 2026

Accepted: 22 June 2026

Published: 25 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Computational Thinking (CT) has become an important focus of computing education, particularly in elementary school contexts where students are increasingly introduced to programming and digital problem-solving activities (Shute et al. 2017; Wing 2006). At the same time, educational inequality continues to shape students' access to digital resources, programming opportunities, and STEM-related learning experiences (Feniger et al. 2021; Yin et al. 2024). These two developments raise an important measurement question: when students from different socioeconomic contexts engage with CT learning, where do such differences become visible?

Despite growing attention to both CT learning and educational inequality, identifying where socioeconomic differences become visible in CT learning remains empirically challenging. One reason is that different assessment approaches capture different aspects of

learning. Diagnostic assessments are designed to capture students' conceptual understanding of computational constructs, whereas digital learning environments generate log-based indicators that document how students act while solving computational tasks, such as their attempts, progression, first-try performance, and time investment (Grover and Pea 2013; Tang et al. 2020). These forms of evidence are related, but they do not necessarily represent the same dimension of CT learning.

The present study addresses this measurement gap by examining socioeconomic differences across three complementary assessment levels: diagnostic CT measures, individual log-based behavioral indicators, and integrated behavioral types derived from patterns of log-based indicators. This design enables us to examine whether socioeconomic variation is visible in conceptual CT knowledge, in isolated behavioral indicators, or in broader configurations of students' digital problem-solving behavior.

The contribution of the study is therefore both substantive and methodological. Substantively, it examines how socioeconomic differences in CT learning are expressed in students' engagement with digital programming tasks. Methodologically, it demonstrates the value of integrating diagnostic and log-based assessment evidence for identifying forms of inequality that may remain hidden when only one assessment modality is used.

Research Questions

The study addresses the following research questions:

1. How are students' diagnostic measures of CT associated with socioeconomic background?
2. How are individual log-based indicators of CT learning associated with socioeconomic background?
3. How are behavioral types, derived from patterns of log-based indicators of CT learning, associated with socioeconomic background?

2. Related Work

2.1. Conceptual and Behavioral Approaches to CT Assessment

Computational Thinking is commonly conceptualized as a multidimensional competence that includes abstraction, decomposition, algorithmic reasoning, debugging, and systematic problem solving (Grover and Pea 2013; Montuori et al. 2024; Palop et al. 2025). As CT has been integrated into primary and secondary education, assessment has become a central concern: researchers and educators need tools that can capture students' conceptual understanding while also recognizing how CT is enacted during problem-solving activity.

Diagnostic CT assessments are designed to evaluate students' conceptual understanding of computational constructs through structured reasoning tasks that are independent of specific learning environments (El-Hamamsy et al. 2025; Tang et al. 2020). These instruments provide standardized measures of conceptual CT knowledge and enable comparisons across students and populations. Validation studies using diagnostic CT assessments, including the Competent Computational Thinking Test (cCTt), have demonstrated their ability to reliably measure students' conceptual understanding of computational constructs in elementary education settings (El-Hamamsy et al. 2025; Zapata-Cáceres et al. 2024).

At the same time, digital programming environments generate detailed log-based process data that capture learners' actions during problem solving, including attempts, progression across tasks, response accuracy, and solution time (Banihashem et al. 2024; He and Cui 2025). These behavioral indicators provide process-level evidence of how learners engage with computational tasks and enact learning opportunities. Empirical studies analyzing log-based interaction data have shown that learners with similar performance outcomes often exhibit distinct behavioral trajectories, including differences in persistence, number of attempts, and progression through programming challenges, indicating that

behavioral data provide insight into learning processes not directly observable through diagnostic assessments alone (Israel-Fishelson and HersHKovitz 2020a; Tan et al. 2025).

Importantly, the multidimensional nature of CT implies that different assessment modalities capture complementary aspects of computational competence. Systematic reviews and empirical studies have emphasized that conceptual understanding and behavioral performance represent related but distinct dimensions of learning, and that no single assessment modality fully captures the complexity of CT development (Montuori et al. 2024). This distinction highlights the importance of integrating multiple assessment approaches to obtain a comprehensive understanding of students' CT.

2.2. Log-Based Behavioral Indicators and Conceptual Knowledge

Learning analytics research has increasingly examined the relationship between behavioral indicators derived from log data and conceptual measures of learning, as documented in recent systematic reviews (Banihashem et al. 2024). Log-based process data provide detailed traces of learners' interactions with digital tasks, capturing behavioral dimensions such as persistence, strategy use, progression, and efficiency during problem solving (Arizmendi et al. 2022; He and Cui 2025). However, empirical research in learning analytics suggests that behavioral trace data and conceptual assessments do not necessarily capture the same aspects of learning (Montuori et al. 2024). Learners who demonstrate similar conceptual understanding may exhibit substantially different behavioral trajectories in digital environments, while similar behavioral patterns do not always correspond to comparable conceptual knowledge.

Studies in digital learning environments show that behavioral trace data may relate to conceptual measures under specific learning conditions and task contexts. Some studies report measurable correspondences between behavioral indicators and conceptual performance when both are derived from the same digital task context (Israel-Fishelson and HersHKovitz 2020b, 2022). For example, empirical analyses of programming and game-based learning environments indicate that indicators such as task completion accuracy, progression patterns, and interaction efficiency may reflect how learners apply domain-specific computational concepts during problem-solving activities within the same instructional setting, rather than constituting direct measures of conceptual understanding. Such convergence appears more likely when behavioral indicators are compared with conceptual assessments that reflect the same learning activity, suggesting that the strength of observed correspondence depends on the degree of contextual alignment between the learning task generating the trace data and the conceptual construct being assessed.

Beyond task-bound comparisons, other studies document weak or inconsistent relationships between behavioral trace data and conceptual assessments administered outside the immediate digital learning context, including in domains such as programming and complex problem solving (Tan et al. 2025; Winter et al. 2024). For instance, learners with similar levels of conceptual performance may exhibit substantial differences in persistence and number of attempts during programming tasks, as well as in their progression patterns and temporal investment in complex problem-solving environments (He and Cui 2025; Tan et al. 2025). These findings suggest that behavioral indicators often capture dimensions of engagement and strategy enactment that are only partially reflected in broader conceptual assessments.

Because learning unfolds through coordinated cognitive and behavioral processes over time, recent research has emphasized the importance of examining integrated behavioral patterns rather than isolated performance indicators (Ferguson et al. 2023). Multidimensional analytical approaches, including clustering methods, enable the identification of behavioral profiles that capture coherent modes of engagement with digital learning envi-

ronments (Ferguson et al. 2023). Recently, latent profile analysis was used to identify CT profiles among pre-service teachers; profile membership was associated with metacognitive strategies and prior coding experience, and behavioral measures from an interactive CT task differed across profiles (Chichekian et al. 2026). Such approaches enable the characterization of systematic variation in how learners engage with computational tasks, thereby providing a process-oriented perspective that complements conceptual assessment (Israel-Fishelson and HersHKovitz 2020a).

2.3. Socioeconomic Differences in CT and Digital Learning

Socioeconomic status (SES) is a well-established predictor of educational outcomes and is consistently associated with students' access to digital resources and learning opportunities across educational contexts (Feniger et al. 2021; Winter et al. 2024; Yin et al. 2024). For example, in primary school settings, students from higher socioeconomic backgrounds achieved significantly higher programming scores and progressed further within the same instructional sequence than their lower-SES peers (Küçükaydın and Çite 2024). Similarly, students with greater home access to computing resources tended to have greater opportunities to engage in programming activities and digital learning experiences (Yin et al. 2024).

Research examining CT specifically has reported mixed findings regarding socioeconomic differences in conceptual CT knowledge. Some studies have found that students from higher-SES backgrounds show higher programming achievement and more advanced performance in programming-related tasks, although this does not necessarily translate into differences in conceptual CT measures, particularly when learners have unequal access to digital learning opportunities (Küçükaydın and Çite 2024). However, other studies have reported weak or statistically non-significant associations between SES and standardized diagnostic measures of conceptual CT, suggesting that SES-related variation is not always detectable at the level of conceptual knowledge alone (Montuori et al. 2024). Research using standardized diagnostic CT assessments emphasizes that such instruments are designed to capture learners' conceptual computational reasoning independently of specific learning contexts or behavioral engagement conditions (El-Hamamsy et al. 2025). This inconsistency suggests that socioeconomic differences in CT may not always be observable through standardized diagnostic conceptual measures.

At the same time, research using log-based interaction data suggests that SES is associated with observable variation in how learners engage with computational tasks. Behavioral trace data analyses have demonstrated higher attempt frequency, slower progression rates, and less stable solution trajectories among lower-SES learners compared to their higher-SES peers (Winter et al. 2024). Because behavioral indicators capture how learners enact learning opportunities, examining behavioral patterns provides an additional perspective for understanding how SES-related variation becomes observable in CT development.

Despite growing research on CT assessment and educational inequality, prior studies have typically examined socioeconomic differences either at the level of conceptual diagnostic assessments or at the level of individual behavioral indicators. However, studies rarely examine conceptual measures, individual behavioral indicators, and integrated behavioral patterns within a unified analytical framework (Banihashem et al. 2024; Ferguson et al. 2023). Without examining these levels simultaneously, interpretations of SES-related variation may remain partial, as each modality foregrounds different dimensions of CT learning. Clarifying where SES-related variation becomes observable across complementary assessment levels is therefore essential for accurately interpreting what different forms of assessment evidence reveal about CT development.

3. Materials and Methods

3.1. Participants and Learning Context

The study involved 444 elementary school students from 3rd and 4th grades (ages 8–10) attending four public elementary schools in a large metropolitan area in Israel. The learning activity focused on the acquisition of core Computational Thinking (CT) concepts and was implemented as part of students' regular classroom instruction, with all students participating in the same instructional activity during regular school hours.

Students worked in the same digital learning environment (CodeMonkey), using a block-based programming interface designed for elementary school students. The learning activity consisted of a structured sequence of goal-oriented programming challenges, in which students constructed solutions by arranging visual code blocks to control a character's actions. The system automatically recorded fine-grained task-level interaction data for each student, documenting how students attempted solutions, progressed across challenges, and invested time during problem solving.

3.2. Research Tools

To examine students' CT from complementary perspectives, the study employed two research tools: a diagnostic pen-and-paper assessment and a digital programming environment that generated fine-grained log data. Together, these tools enabled the examination of students' conceptual CT knowledge alongside their observed problem-solving behavior in a digital learning context.

3.2.1. Competent Computational Thinking Test (cCTt)

Students' conceptual CT knowledge was assessed using the Competent Computational Thinking Test (cCTt), a validated pen-and-paper diagnostic instrument designed for elementary school students (El-Hamamsy et al. 2025). The cCTt evaluates students' understanding of core computational concepts through multiple-choice items that require reasoning about algorithmic processes rather than code writing. In the present study, we used a validated shortened version of the instrument consisting of 15 multiple-choice items, selected from the original 25-item test and organized to represent core computational concepts with progressive complexity.

The test targets foundational CT concepts commonly taught in introductory programming instruction, including sequences, loops, conditionals, and while statements. Items are presented in abstract, non-programming contexts and require students to interpret, predict, or reason about the behavior of computational processes. The cCTt yields both concept-level scores and an overall CT performance score, providing an estimate of students' diagnostic conceptual CT knowledge independent of interaction with a specific programming environment. The cCTt has demonstrated satisfactory psychometric properties in prior studies with elementary school populations and has been used in multiple large-scale investigations of CT development (El-Hamamsy et al. 2025). Illustrative examples of cCTt items are provided in Figure 1. The examples present abstract grid-based representations of a character's movement in a multiple-choice format with four response options. Item 1 represents a sequencing task, whereas Item 13 represents a task involving repeated actions.

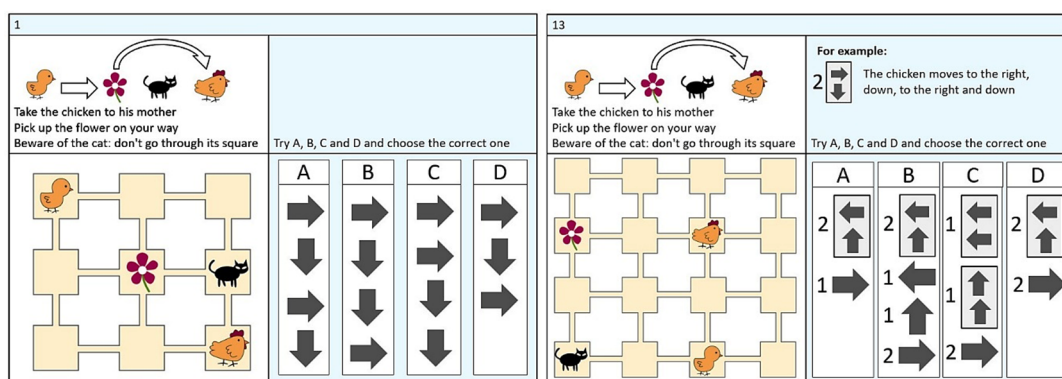


Figure 1. Sample items from the cCTt instrument. Item 1 shows a sequencing task, and Item 13 shows a task that involves repeated actions. Arrows indicate movement directions, the numbers 1 and 13 identify the item numbers, and the letters A–D label the response options. The figure illustrates the grid-based problem context and the four-option multiple-choice format.

3.2.2. CodeMonkey Digital Learning Environment

Students' digital learning behavior was captured through CodeMonkey, a block-based programming environment designed for elementary school learners. In CodeMonkey, students engage with a structured and progressive sequence of programming challenges of increasing complexity by constructing solutions through the arrangement of visual code blocks that control a character's actions.

Throughout students' interactions with the environment, CodeMonkey automatically records time-stamped task-level log data that capture individual interaction events, including solution attempts, challenge completion, and time spent on each challenge. Behavioral indicators used in this study were derived from these raw log records through data processing and aggregation procedures. These log data provide detailed information about students' engagement and interaction patterns within the digital learning environment and serve as the basis for the log-based analyses conducted in this study. Figure 2 shows examples of CodeMonkey challenges from the Beaver Achiever unit. The top panel shows an early challenge, whereas the bottom panel shows a later challenge that involves organizing repeated actions using loop blocks. The examples illustrate the block-based programming interface and task context.

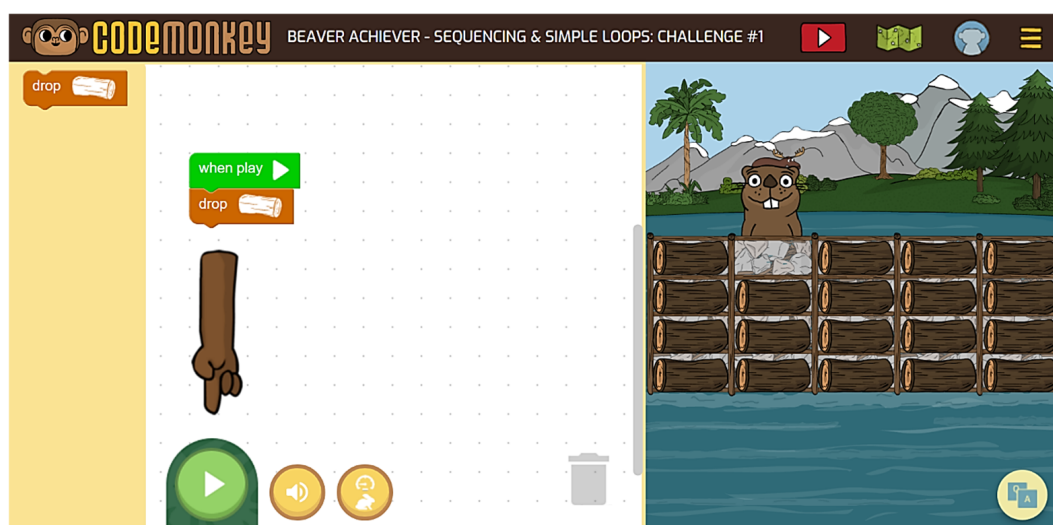


Figure 2. Cont.

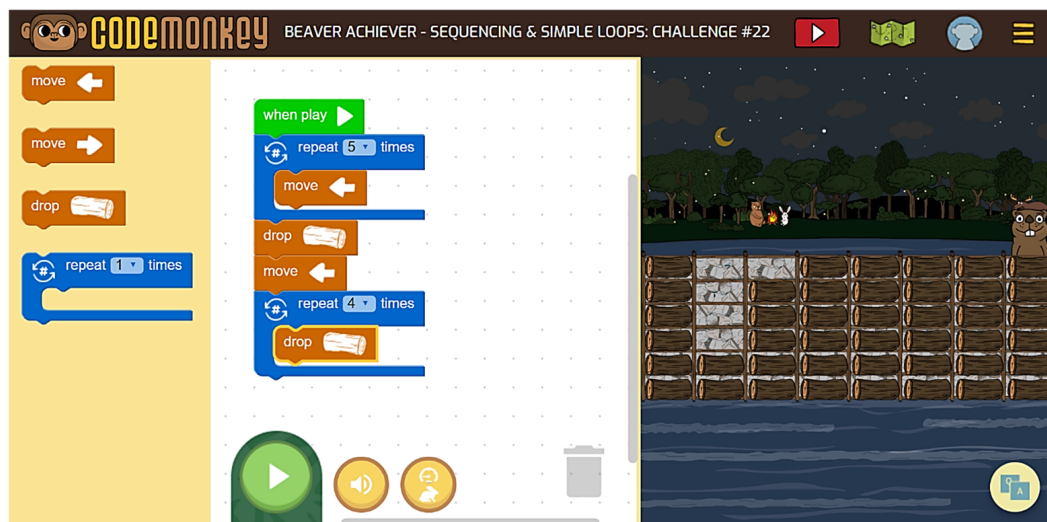


Figure 2. Examples of programming challenges from the CodeMonkey Beaver Achiever unit. Top panel shows Challenge 1. Bottom panel shows Challenge 22. The # symbol in the repeat block indicates the number of times the enclosed command sequence is repeated. The figure illustrates the block-based programming interface and task context.

3.3. Measures

To capture students' CT from complementary perspectives, the study included diagnostic CT measures, log-based behavioral measures derived from students' interaction data, and a school-level socioeconomic indicator.

3.3.1. Diagnostic CT Measures

Students' diagnostic CT was operationalized using scores derived from the cCTt. The test yields both concept-level scores (sequences, loops, conditionals, and while statements) and an overall CT performance score, which were used as indicators of students' conceptual CT knowledge independent of their interaction with the digital learning environment.

3.3.2. Log-Based Behavioral Measures

Log-based behavioral measures were derived from students' interactions with the digital learning environment and treated as process-level indicators reflecting how students acted while engaging with CT tasks. As students progressed through a structured sequence of programming challenges, the system recorded detailed logs of their problem-solving activity, capturing first-try performance, persistence, progression, and time investment.

Based on these data, four behavioral indicators were constructed: average first-try stars, defined as the mean number of stars obtained on students' first attempt across completed challenges; attempts per challenge, reflecting the average number of solution attempts and capturing persistence and trial-and-error behavior; highest challenge reached, indicating students' progression through the task sequence; and average time per challenge, representing temporal investment during problem solving. Together, these indicators provide a multidimensional characterization of students' digital problem-solving behavior.

These indicators were selected because they represent interpretable dimensions of digital learning behavior commonly examined in current learning analytics research: initial task performance, repeated attempts and persistence, progression through a structured task sequence, and temporal investment during problem solving (Arizmendi et al. 2022; Banihashem et al. 2024; He and Cui 2025). Although this feature set does not model fine-grained temporal sequences or strategy transitions, it provides a parsimonious basis for comparing broad log-based behavioral evidence with diagnostic CT performance.

3.3.3. Background Variables

Socioeconomic status (SES) was operationalized at the school level using the Israeli Ministry of Education's Nurture Index (BenDavid-Hadar and Ziderman 2011), with lower values indicating higher socioeconomic status. This indicator was used to examine socioeconomic differences across diagnostic CT measures and log-based behavioral outcomes. Because the available SES indicator was defined at the school level, SES was treated as a contextual school-level grouping variable rather than as an individual-level socioeconomic measure. Student-level background variables were not included as covariates in the main analyses; this analytic choice is addressed as a limitation.

3.4. Data Analysis

Analyses were conducted at two complementary levels. First, socioeconomic differences were examined separately for diagnostic CT scores derived from the cCTt and for individual log-based behavioral indicators derived from the digital learning environment.

Second, students' digital behaviors were analyzed as structured patterns of action. Log-based indicators were z-standardized and subjected to hierarchical clustering using Ward's linkage method with Pearson distance to identify digital behavioral types based on students' interaction patterns. Cluster solutions were evaluated using a combination of BIC, silhouette values, the elbow plot, dendrogram inspection, cluster size, and substantive interpretability.

All analyses were performed using JASP (version 0.95). Missing values were handled using analysis-specific deletion, so the analytic sample size varied across analyses according to the availability of the relevant diagnostic, log-based, and clustering variables.

4. Results

4.1. Diagnostic Measures and Socioeconomic Status (RQ1)

To examine socioeconomic differences in students' diagnostic CT performance, participants were grouped based on school-level SES. The four participating schools were mapped into two SES groups, representing lower-SES and higher-SES school contexts. Among students with valid school-level SES information, the higher-SES group comprised 104 students, while the lower-SES group comprised 263 students. All subsequent analyses were conducted at the student level, using this dichotomous SES grouping as a contextual variable rather than as an analytic unit at the school level.

Descriptive statistics were computed for cCTt scores separately for each SES group. Due to missing cCTt data, this analysis includes 101 students from higher-SES schools and 260 students from lower-SES schools. Students in the higher-SES group obtained a mean cCTt score of $M = 0.66$ ($SD = 0.28$), whereas students in the lower-SES group obtained a mean score of $M = 0.67$ ($SD = 0.25$). An independent-samples *t*-test indicated no statistically significant difference between groups, $t(359) = -0.18$, $p = 0.86$.

4.2. Individual Log-Based Indicators (RQ2)

Socioeconomic differences in individual log-based indicators were examined by comparing students from lower- and higher-SES school contexts across four predefined behavioral measures: average first-try stars, average attempts per challenge, highest challenge reached, and average time per challenge. For each indicator, an independent-samples *t*-test was conducted to assess group differences.

Across the individual log-based indicators, most socioeconomic differences were not statistically significant. No significant differences between lower- and higher-SES school contexts were found for the average number of attempts per challenge, $t(371) = -1.13$, $p = 0.26$, the highest challenge reached, $t(371) = -1.44$, $p = 0.15$, or the average time per

challenge, $t(371) = 1.90$, $p = 0.06$. A statistically significant difference between groups was observed only for average first-try stars, $t(371) = 5.57$, $p < 0.001$, with students from higher-SES school contexts demonstrating higher average first-try stars. Descriptive statistics and results of the independent-samples t -tests for all individual log-based indicators are presented in Table 1. Overall, socioeconomic differences were observed for only one of the examined log-based indicators and were not consistently present across individual behavioral measures.

Table 1. Descriptive Statistics and SES Comparisons for Individual Log-Based Indicators.

Log-Based Indicator	High-SES (n = 95) M (SD)	Low-SES (n = 278) M (SD)	t	p
Average first-try stars	2.64 (0.27)	2.42 (0.34)	5.57	<0.001
Average attempts per challenge	1.51 (0.46)	1.58 (0.54)	−1.13	0.26
Highest challenge reached	23.23 (4.74)	24.07 (4.94)	−1.44	0.15
Average time per challenge (seconds)	86.48 (65.98)	76.80 (31.50)	1.90	0.06

Note. SES = socioeconomic status. Values are means (M) and standard deviations (SD) by school-level SES group. Group differences were tested using independent-samples t -tests. Reported n values reflect valid (non-missing) observations for each indicator.

4.3. Digital Behavioral Types (RQ3)

To address the third research question, the analysis moved beyond examining individual log-based indicators in isolation and instead focused on identifying integrated patterns of digital learning behavior. Rather than analyzing each behavioral measure separately, this step aimed to capture how multiple indicators jointly characterize students' interaction with the CodeMonkey learning environment. Given the multidimensional nature of log-based data, hierarchical clustering was applied to standardized measures of average first-try stars, average attempts per challenge, highest challenge reached, and average solution time.

Hierarchical clustering identified four distinct behavioral clusters based on students' log-based interaction patterns. The clustering solution was estimated using Pearson distance and Ward's linkage method. Model fit indices indicated meaningful separation between clusters, with an explained variance of $R^2 = 0.45$ and an average silhouette coefficient of 0.58. The four-cluster solution was selected as the primary model based on the balance between BIC, silhouette values, the elbow plot, dendrogram inspection, cluster size, and substantive interpretability, whereas higher-k solutions yielded smaller clusters without adding substantive interpretive value.

To characterize the behavioral patterns represented by each cluster, cluster-level means were computed for the four standardized log-based indicators, enabling direct comparison of standardized indicator values across clusters. Rather than reflecting variation along a single behavioral dimension, the clusters differ in the configuration and co-occurrence of multiple indicators, as reflected in their standardized values across the four behavioral indicators. Table 2 presents the standardized profiles of the four behavioral types across average first-try stars, average attempts per challenge, highest challenge reached, and average solution time, alongside the corresponding cCTt means and SES distribution. Figure 3 presents a t-SNE visualization of the four-cluster solution based on students' standardized log-based behavioral indicators, showing the separation between clusters in a reduced two-dimensional space. Differences in overall cCTt scores across behavioral types were examined using a one-way ANOVA. The analysis revealed statistically significant differences between clusters, $F(3, 321) = 16.30$, $p < 0.001$. Mean cCTt scores were 0.52 (SD = 0.25) for Cluster 1, 0.66 (SD = 0.26) for Cluster 2, 0.63 (SD = 0.26) for Cluster 3, and 0.78 (SD = 0.21) for Cluster 4.

Table 2. Behavioral Cluster Characteristics and Distribution by SES School Context (k = 4).

Cluster	N	Avg. First-Try Stars	Avg. Attempts per Challenge	Highest Challenge Reached	Avg. Solution Time	cCTt Mean	cCTt SD	% High-SES
Cluster 1 (High Attempts, Low Progression)	74	−0.30	1.25	−0.98	0.26	0.52	0.25	21.62
Cluster 2 (High First-Try Stars, Slow, Low Progression)	85	0.57	−0.07	−0.85	0.58	0.66	0.26	38.82
Cluster 3 (Low First-Try Stars, Fast Progression)	92	−1.21	−0.28	0.57	−0.30	0.63	0.26	9.78
Cluster 4 (High First-Try Stars, Efficient Progression)	122	0.70	−0.50	0.76	−0.33	0.78	0.21	30.33

Note. Log-based indicators are reported as standardized cluster means (z-scores). SES distribution is reported as row percentages. cCTt values represent cluster-level means and standard deviations of overall cCTt scores.

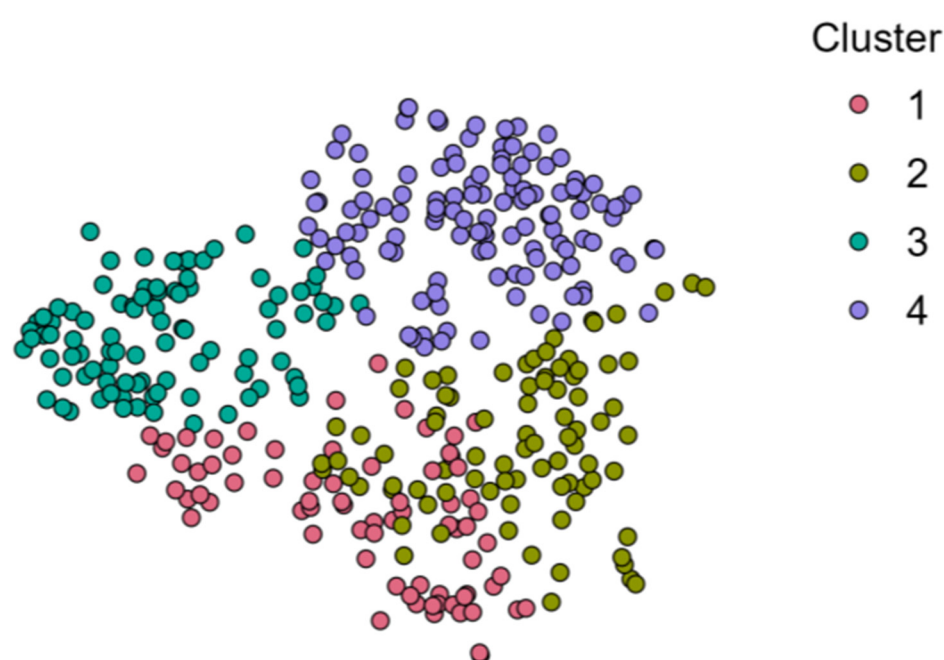


Figure 3. t-SNE visualization of the four behavioral clusters based on standardized log-based indicators. Each point represents a student, and colors denote cluster membership identified through hierarchical clustering.

Cluster-level values for the log-based indicators represent standardized scores (z-scores). Values of 0 indicate the sample average, whereas positive and negative values indicate above- and below-average performance, respectively. These standardized profiles provide the basis for interpreting the behavioral patterns represented by each cluster.

Cluster 1 was characterized by below-average first-try performance, substantially above-average attempts per challenge, substantially below-average highest challenge reached, and above-average solution time. Based on this standardized profile, this behavioral type is labeled High Attempts, Low Progression.

Cluster 2 showed above-average first-try performance, near-average attempts per challenge, below-average highest challenge reached, and substantially above-average solution time. Based on this standardized profile, this behavioral type is labeled High First-Try Stars, Slow, Low Progression.

Cluster 3 was characterized by substantially below-average first-try performance, below-average attempts per challenge, above-average highest challenge reached, and

below-average solution time. Based on this standardized profile, this behavioral type is labeled Low First-Try Stars, Fast Progression.

Cluster 4 showed above-average first-try performance, below-average attempts per challenge, above-average highest challenge reached, and below-average solution time. Based on this standardized profile, this behavioral type is labeled High First-Try Stars, Efficient Progression.

After establishing the behavioral cluster structure, a chi-square test of independence was conducted to examine whether cluster membership was associated with school-level SES context. The analysis indicated a statistically significant association between behavioral type and SES group, $\chi^2(3) = 22.01, p < 0.001$ ($N = 373$), with a small-to-moderate effect size (Cramer's $V = 0.24$). The SES distribution across clusters showed differential representation of SES groups across specific behavioral profiles. Students from higher-SES schools comprised 25.5% of the clustering sample but accounted for 38.8% of Cluster 2 and 30.3% of Cluster 4. In contrast, students from lower-SES schools comprised 74.5% of the clustering sample but accounted for 90.2% of Cluster 3.

To complement the tabular presentation in Table 2 and the reduced-space visualization in Figure 3, Figure 4 presents the standardized behavioral profiles of the four clusters and the SES distribution within each behavioral type. This visualization makes the contrast between assessment levels more explicit: diagnostic CT scores did not differ significantly between SES groups, whereas SES-related differences became visible in the distribution of students across multidimensional behavioral profiles.

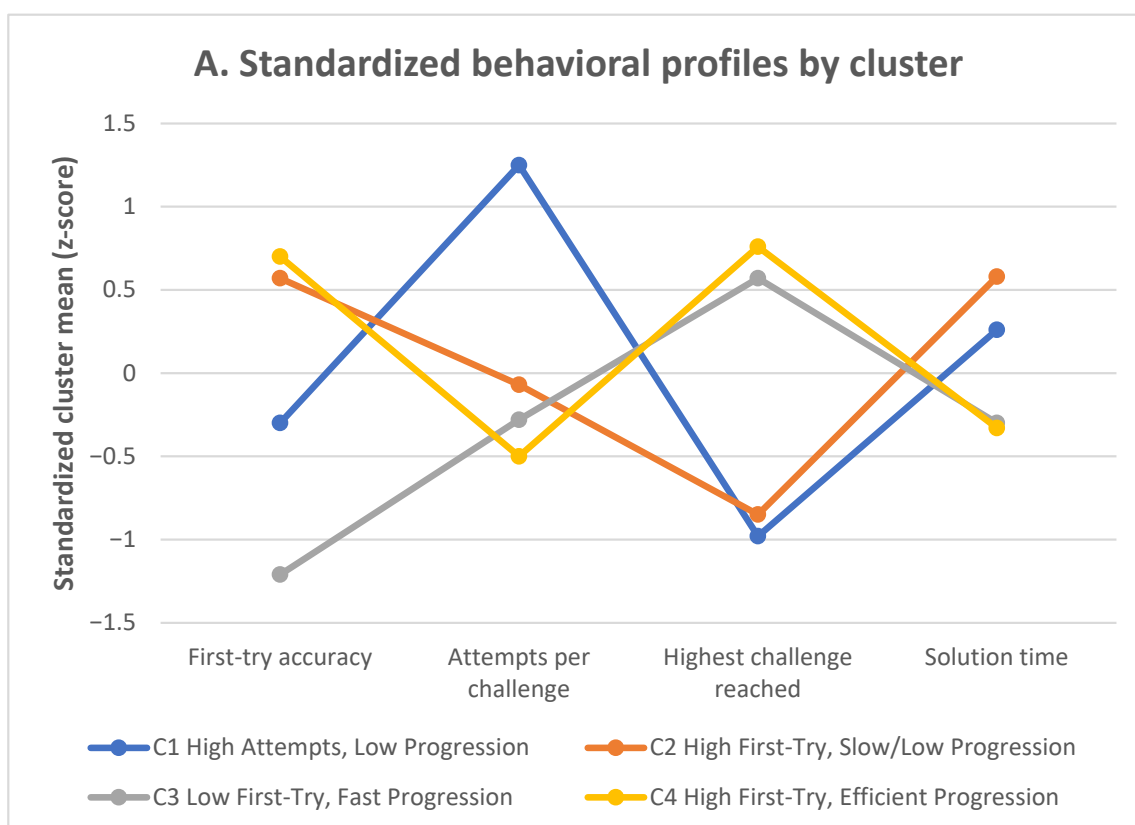


Figure 4. Cont.

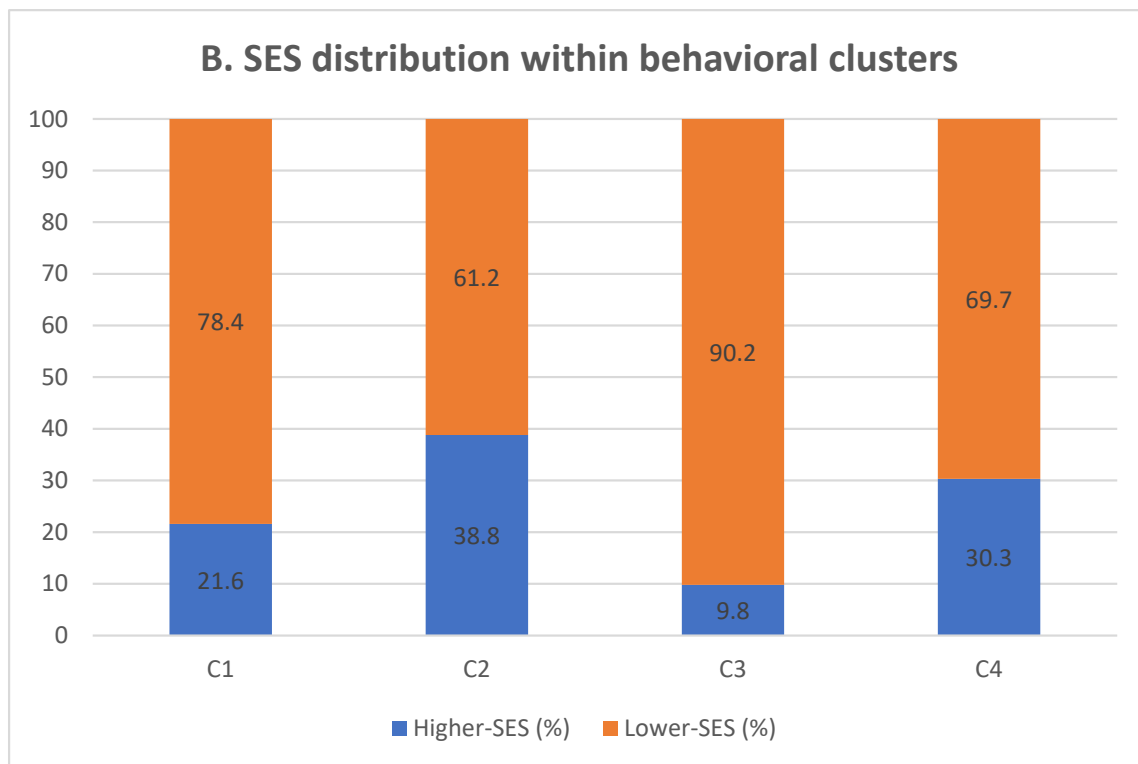


Figure 4. Behavioral profiles and SES distribution across behavioral types. Panel (A) presents the standardized cluster means for the four log-based behavioral indicators. Panel (B) presents the percentage of higher- and lower-SES students within each behavioral type. Higher values in Panel A indicate values above the sample mean; for attempts and solution time, higher values indicate more attempts and longer solution time. Diagnostic CT scores did not differ significantly by SES group, whereas SES-related differences became visible in the distribution of students across multidimensional behavioral profiles.

5. Discussion

5.1. Differences in CT Measures Across Socioeconomic Groups

In the present study, socioeconomic differences were largely absent when CT was examined through individual indicators derived from either the diagnostic assessment or the programming environment. This was observed across most log-based behavioral indicators, including attempts per challenge, progression, and solution time. The only indicator that differentiated between socioeconomic groups was average first-try stars. Notably, a similar absence of socioeconomic differences was observed in the diagnostic assessment, suggesting that CT performance, as captured by the diagnostic measure, may be less sensitive to SES-related variation in how learners enact computational tasks.

These findings should be interpreted in light of recent research showing that socioeconomic differences in CT and digital learning are shaped not only by students' final achievement, but also by differences in access to digital resources, programming opportunities, school-level technology use, and broader conditions of digital participation (Gottschalk and Weise 2023; Küçükaydın and Çite 2024). Thus, the absence of SES differences in diagnostic CT scores does not necessarily indicate the absence of inequality; rather, it suggests that inequality may be more visible in how students engage with digital programming tasks. From a measurement perspective, the present study addresses a current gap in CT assessment: recent studies show that outcome-focused CT assessments may overlook process-level, cognitive, and behavioral differences that become visible when learners' task processes and multimodal evidence are examined (Bhatt et al. 2024; Yang et al. 2025).

By comparing diagnostic CT scores with individual and integrated log-based indicators within the same learning context, the present study shows that the assessment method itself shapes where SES-related variation becomes observable.

Regarding the single measure that was found to demonstrate SES-based differences, i.e., average first-try stars, research on programming learning environments suggests that learners' approaches to programming tasks may be shaped by differences in prior exposure to programming opportunities and digital learning experiences (Grover and Pea 2013; Yin et al. 2024). Access to such experiences is demographically biased, as students from higher socioeconomic backgrounds are more likely to encounter computing technologies and programming opportunities outside formal schooling (Küçükaydın and Çite 2024). Under these conditions, learners may demonstrate comparable diagnostic CT performance and similar persistence during iterative problem solving, while still differing in their patterns of performance when first encountering a task.

From a measurement perspective, this distinction suggests that indicators capturing immediate task success may be particularly sensitive to differences in learners' readiness at the point of initial task engagement (Yin et al. 2024). In contrast, persistence-oriented indicators such as attempts per challenge or time investment reflect learners' iterative problem-solving behavior during task engagement rather than immediate task readiness (Israel-Fishelson and HersHKovitz 2020b, 2022). These differences reflect the multi-faceted nature of log-based CT indicators, which capture distinct dimensions of how learners enact computational knowledge during task interaction. Similar patterns have been reported in digital learning research, where indicators reflecting early task performance revealed differences associated with prior learning opportunities even when persistence-related measures did not (Banihashem et al. 2024).

Taken together, these findings suggest that socioeconomic variation in CT may not always be detectable through isolated indicators. Instead, such variation may emerge through the joint organization of multiple behavioral dimensions during interaction with computational tasks. Learning analytics research has therefore emphasized the value of person-centered approaches for identifying latent behavioral patterns that remain hidden in variable-centered analyses (Dijkstra et al. 2023; Saqr 2023). These findings motivate the transition to pattern-level analyses examining whether socioeconomic variation becomes more visible when learning processes are modeled as multidimensional configurations.

5.2. Identification of Digital Behavioral Types

In the present study, hierarchical clustering of multiple log-based indicators revealed distinct behavioral configurations in how students engaged with the digital programming environment. This interpretation is consistent with learning analytics research showing that digital interaction data often reflect coordinated aspects of learner activity and that person-centered and multivariate analytical approaches can identify empirically interpretable profiles of engagement across learning tasks (Arizmendi et al. 2022; Banihashem et al. 2024; Dijkstra et al. 2023). Accordingly, the clusters are interpreted here as context-specific behavioral configurations observed within the digital programming task environment, rather than as stable learner types or externally validated learner categories.

Prior research in learning analytics has shown that log-based interaction data often contain latent behavioral structures that become visible when multiple indicators are analyzed together. Empirical studies using clustering and person-centered analytical approaches demonstrate that indicators such as persistence, efficiency, and progression can form distinct configurations across learners because learning processes unfold through coordinated behavioral and cognitive actions over time. As a result, multivariate analyses reveal structured engagement profiles that reflect how learners balance multiple demands during task

interaction, rather than isolated performance features (Arizmendi et al. 2022; Dijkstra et al. 2023). Systematic reviews further indicate that such multivariate behavioral configurations provide more robust representations of engagement by capturing interdependencies between complementary dimensions of activity (Banihashem et al. 2024). This perspective provides a theoretical basis for interpreting the behavioral configurations identified in the present study as meaningful representations of how learners engage with computational problem-solving tasks, rather than as methodological artifacts of the clustering procedure.

In the present study, the identified behavioral types reflected distinct configurations of engagement across multiple indicators, indicating that students approached the same sequence of computational tasks through different coordinated patterns of action. The identification of behavioral types thus supports a complex interpretation of digital engagement, in which variation in students' engagement is understood through coordinated configurations of activity rather than through isolated indicators (Banihashem et al. 2024).

5.3. Socioeconomic Differences in Digital Behavioral Types

Building on the first two findings, socioeconomic differences became visible when students' activity in the digital learning environment was examined through the behavioral types identified in the clustering analysis. Socioeconomic differences were reflected in the uneven representation of SES groups across behavioral types characterized by distinct configurations of first-try performance, persistence, progression, and temporal investment. Students from higher-SES school contexts were relatively more represented in clusters characterized by above-average first-try performance, although these clusters differed in progression and time investment. In contrast, students from lower-SES school contexts were especially overrepresented in the cluster characterized by low first-try performance combined with relatively fast progression, and were also slightly overrepresented in the high-attempts, low-progression cluster. Thus, the SES-related pattern did not reflect a simple high-versus-low performance gradient, but rather different configurations of how students engaged with the digital learning environment.

Related divergences between conceptual performance and behavioral trajectories have been documented in research on digital learning and complex problem solving, where learners with comparable diagnostic performance exhibit different interaction patterns reflecting variation in uncertainty management and strategy enactment (Tan et al. 2025; Winter et al. 2024). These patterns suggest that socioeconomic differences were manifested primarily in how learners progressed, persisted, and invested time while engaging with programming challenges, rather than in their diagnostic CT performance.

A possible interpretation of this pattern relates to differences in prior exposure to structured programming learning environments and in the regulation of exploratory behavior during problem solving. Learners with greater prior exposure may progress more rapidly through programming sequences by adopting risk-taking or exploration strategies, even when first-try performance is lower. In contrast, learners with limited exposure may adopt more cautious or time-intensive approaches, resulting in slower progression or greater reliance on repeated attempts (Brender et al. 2025; Schneider and Simonsmeier 2025). An alternative, complementary interpretation is that such differences may also reflect variation in how learners regulate uncertainty and strategic decision-making during problem solving, rather than prior exposure alone (Bohm et al. 2024).

5.4. Implications and Limitations

The findings suggest several implications for research, instructional practice, and the design of digital learning environments. For researchers, the results indicate that socioeconomic variation may remain undetected when analyses rely on a single assessment

modality, such as diagnostic measures or isolated behavioral indicators. Future research should therefore integrate diagnostic and log-based measures to more accurately capture how inequality is enacted during digital problem-solving processes. For educators, the findings suggest that students who demonstrate similar levels of diagnostic CT performance may nevertheless differ in how they progress, persist, and invest time when working with digital programming tasks. Instructional practices should thus account for differences in learning processes, not only for differences in conceptual performance. For developers, the results suggest that analytic tools should be designed that capture coordinated patterns of learner activity rather than focusing only on single performance metrics.

At the same time, several limitations should be acknowledged. First, the study was conducted within one digital programming environment, which may limit the generalizability of the behavioral patterns identified. In addition, the study involved a limited population of 3rd- and 4th-grade students from four public elementary schools in one metropolitan region in Israel, and the findings should therefore be interpreted with caution when considering other age groups, educational systems, or national contexts. Second, the behavioral indicators examined in this study represent a limited set of log-derived behavioral measures and therefore do not capture the full range of cognitive and contextual factors that may shape students' engagement with computational tasks. Third, socioeconomic status was operationalized using the Ministry of Education's school-level Nurture Index rather than individual-level socioeconomic data. This provided a contextual indicator of students' learning environments, but it does not fully represent the complexity of students' individual socioeconomic backgrounds or the broader socioeconomic influences that may shape learning.

In addition, the analyses did not include individual-level covariates such as gender, prior programming experience, or cognitive ability. Therefore, the findings should be interpreted as associations between school-level SES context and assessment patterns, rather than as evidence of causal SES effects. Finally, the log-based indicators used in this study provide a parsimonious summary of students' digital behavior but do not capture fine-grained temporal dynamics, sequential strategies, or longitudinal development. Future research should therefore combine diagnostic assessments, richer temporal learning analytics, individual-level socioeconomic measures, and mixed-methods approaches.

6. Conclusions

This study examined where socioeconomic differences in CT become visible when comparing diagnostic assessment results with behavioral indicators derived from a digital learning environment. The findings show that socioeconomic differences did not manifest in students' diagnostic CT scores and were largely absent across individual log-based performance indicators. However, such differences became visible when students' interaction data were examined as integrated behavioral patterns.

Taken together, these findings indicate that socioeconomic differences in CT learning become visible primarily in how students engage with digital learning environments rather than in diagnostic measures or most isolated behavioral indicators. By capturing coordinated patterns of first-try performance, persistence, progression, and temporal investment, multidimensional analyses of interaction data provide a more sensitive perspective for understanding where and how socioeconomic differences emerge in digital programming contexts.

Author Contributions: Conceptualization, B.A.-L. and A.H.; methodology, B.A.-L. and A.H.; formal analysis, B.A.-L.; investigation, B.A.-L.; data curation, B.A.-L.; writing—original draft preparation, B.A.-L.; writing—review and editing, B.A.-L. and A.H.; visualization, B.A.-L.; supervision,

A.H.; project administration, B.A.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by scholarships awarded to Ben Avital-Lev by the School of Education at Tel Aviv University and the Israel Scholarship Education Foundation (ISEF). No grant numbers are applicable. The article processing charge (APC) was not externally funded.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Tel Aviv University; (Approval No. 0010674-2) and by the Israeli Ministry of Education (Approval No. 14090).

Informed Consent Statement: Parental informed consent was obtained through an opt-out procedure approved by the relevant authorities. Parents or legal guardians were informed in advance, and all data were anonymized prior to analysis.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy and ethical restrictions related to research involving minors.

Acknowledgments: The authors thank the participating schools, teachers, and students for their collaboration and support throughout the study. We also thank the CodeMonkey team for enabling us to use their platform and for their assistance with data curation.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Arizmendi, Cara J., Matthew L. Bernacki, Mladen Raković, Robert D. Plumley, Christopher J. Urban, Abigail T. Panter, Jeffrey A. Greene, and Kathleen M. Gates. 2022. Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods* 55: 3026–54. [\[CrossRef\]](#) [\[PubMed\]](#)
- Banihashem, Seyyed Kazem, Hojjat Dehghanzadeh, Douglas Clark, Omid Noroozi, and Harm J. A. Biemans. 2024. Learning analytics for online game-based learning: A systematic literature review. *Behaviour & Information Technology* 43: 2689–716. [\[CrossRef\]](#)
- BenDavid-Hadar, Iris, and Adrian Zideman. 2011. A new model for equitable and efficient resource allocation to schools: The Israeli case. *Education Economics* 19: 341–62. [\[CrossRef\]](#)
- Bhatt, Sohumi M., Katrien Verbert, and Wim Van Den Noortgate. 2024. A method for developing process-based assessments for computational thinking tasks. *Journal of Learning Analytics* 11: 157–73. [\[CrossRef\]](#)
- Bohm, Nina L., Renate G. Klaassen, Ellen van Bueren, and Perry den Brok. 2024. How do students deal with the uncertainty of sustainability challenges? Metacognitive learning in a transdisciplinary course. *Frontiers in Education* 9: 1362075. [\[CrossRef\]](#)
- Brender, Jérôme, Laila El-Hamamsy, Christian Giang, Laura Mathex, Tanja Käser, and Francesco Mondada. 2025. The handwriting programming language for primary school: Aligning computer science education with established teaching practices. *Educational Technology Research and Development* 73: 2699–738. [\[CrossRef\]](#)
- Chichekian, Tanya, Maria Cutumisu, Annie Savard, and Yi-Mei Zhang. 2026. Profiling pre-service teachers' computational thinking: The role of metacognition and coding experience. *Journal of Learning Analytics* 13: 31–41. [\[CrossRef\]](#)
- Dijkstra, Rick, Max Hinne, Eliane Segers, and Inge Molenaar. 2023. Clustering children's learning behaviour to identify self-regulated learning support needs. *Computers in Human Behavior* 145: 107754. [\[CrossRef\]](#)
- El-Hamamsy, Laila, María Zapata-Cáceres, Estefanía Martín-Barroso, Francesco Mondada, Jessica Dehler Zufferey, Barbara Bruno, and Marcos Román-González. 2025. The Competent Computational Thinking Test (cCTt): A Valid, Reliable and Gender-Fair Test for Longitudinal CT Studies in Grades 3–6. *Technology, Knowledge and Learning* 30: 1607–61. [\[CrossRef\]](#)
- Feniger, Yariv, Yossi Shavit, and Shir Caller. 2021. Schooling and Equity in Israel. *The Oxford Encyclopedia of Race and Education* 1: 679–93. [\[CrossRef\]](#)
- Ferguson, Rebecca, Hassan Khosravi, Vitomir Kovanović, Olga Viberg, Ashish Aggarwal, Matthieu Brinkhuis, Simon Buckingham Shum, Lujie Karen Chen, Hendrik Drachler, Valerie A. Guerrero, and et al. 2023. Aligning the goals of learning analytics with its research scholarship: An open peer commentary approach. *Journal of Learning Analytics* 10: 14–50. [\[CrossRef\]](#)
- Gottschalk, Francesca, and Crystal Weise. 2023. Digital equity and inclusion in education: An overview of practice and policy in OECD countries. In *OECD Publishing*. OECD Education Working Papers. Paris: OECD, vol. 299. [\[CrossRef\]](#)

- Grover, Shuchi, and Roy D. Pea. 2013. Computational thinking in K–12: A review of the state of the field. *Educational Researcher* 42: 38–43. [\[CrossRef\]](#)
- He, Surina, and Ying Cui. 2025. A systematic review of the use of log-based process data in computer-based assessments. *Computers & Education* 228: 105245. [\[CrossRef\]](#)
- Israel-Fishelson, Rotem, and Arnon HersHKovitz. 2020a. Persistence in a Game-Based Learning Environment: The Case of Elementary School Students Learning Computational Thinking. *Journal of Educational Computing Research* 58: 891–918. [\[CrossRef\]](#)
- Israel-Fishelson, Rotem, and Arnon HersHKovitz. 2020b. Shooting for the stars: Micro-persistence of students in game-based learning environments. In *Early Warning Systems and Targeted Interventions for Student Success in Online Courses*. Hershey: IGI Global, pp. 239–58. [\[CrossRef\]](#)
- Israel-Fishelson, Rotem, and Arnon HersHKovitz. 2022. Log-based multidimensional measurement of CT acquisition. Paper presented at the Sixth APSCE International Conference on Computational Thinking and STEM Education (CTE-STEM 2022), Delft, The Netherlands, June 15–17; pp. 128–33. [\[CrossRef\]](#)
- Küçükaydın, Menşure Alkış, and Hakan Çite. 2024. Computational thinking in primary school: Effects of student and school characteristics. *Education and Information Technologies* 29: 5631–49. [\[CrossRef\]](#)
- Montuori, Chiara, Filippo Gambarota, Gianmarco Altoé, and Barbara Arfé. 2024. The cognitive effects of computational thinking: A systematic review and meta-analytic study. *Computers & Education* 210: 104961. [\[CrossRef\]](#)
- Palop, Belén, Irene Díaz, Luis J. Rodríguez-Muñiz, and Juan José Santaengracia. 2025. Redefining computational thinking: A holistic framework and its implications for K-12 education. *Education and Information Technologies* 30: 13385–410. [\[CrossRef\]](#)
- Saqr, Mohammed. 2023. Modelling within-person idiographic variance could help explain and individualize learning. *British Journal of Educational Technology* 54: 1077–94. [\[CrossRef\]](#)
- Schneider, Michael, and Bianca A. Simonsmeier. 2025. How does prior knowledge affect learning? A review of 16 mechanisms and a framework for future research. *Learning and Individual Differences* 122: 102744. [\[CrossRef\]](#)
- Shute, Valerie J., Chen Sun, and Jodi Asbell-Clarke. 2017. Demystifying computational thinking. *Educational Research Review* 22: 142–58. [\[CrossRef\]](#)
- Tan, Lihua, Bing Wei, and Fu Chen. 2025. An exploratory process mining on students' complex problem-solving behavior: The distinct patterns and related factors. *Computers & Education* 238: 105398. [\[CrossRef\]](#)
- Tang, Xiaodan, Yue Yin, Qiao Lin, Roxana Hadad, and Xiaoming Zhai. 2020. Assessing computational thinking: A systematic review of empirical studies. *Computers & Education* 148: 103798. [\[CrossRef\]](#)
- Wing, Jeannette M. 2006. Computational thinking. *Communications of the ACM* 49: 33–35. [\[CrossRef\]](#)
- Winter, Marc, Julia Mordel, Julia Mendzheritskaya, Daniel Biedermann, George-Petru Ciordas-Hertel, Carolin Hahnel, Daniel Bengs, Ilka Wolter, Frank Goldhammer, Hendrik Drachsler, and et al. 2024. Behavioral trace data in an online learning environment as indicators of learning engagement in university students. *Frontiers in Psychology* 15: 1396881. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yang, Sa, Sha Zhu, and Wei Qin. 2025. Enhancing computational thinking assessment: A multimodal cognitive diagnostic approach. *Thinking Skills and Creativity* 56: 101752. [\[CrossRef\]](#)
- Yin, Stella Xin, Dion Hoe-Lian Goh, and Choon Lang Quek. 2024. Collaborative Learning in K-12 Computational Thinking Education: A Systematic Review. *Journal of Educational Computing Research* 62: 1440–74. [\[CrossRef\]](#)
- Zapata-Cáceres, María, Pedro Marcelino, Laila El-Hamamsy, and Estefanía Martín-Barroso. 2024. A Bebras Computational Thinking (ABC-Thinking) program for primary school: Evaluation using the competent computational thinking test. *Education and Information Technologies* 29: 14969–98. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.