Received: 5 April 2023 | Revised: 13 June 2023 | Accepted: 29 June 2024
DOI: 10.1111/jcal.13041

ORIGINAL ARTICLE

Journal of Computer Assisted Learning WILEY

Educational policy as predictor of computational thinking: A supervised machine learning approach

Ndudi O. Ezeamuzie[1,2] | Jessica S. C. Leung[1] | Dennis C. L. Fung[1] | Mercy N. Ezeamuzie[3]

[1]Faculty of Education, University of Hong Kong, Pokfulam, Hong Kong
[2]School of Science and Technology, Hong Kong Metropolitan University, Kowloong, Hong Kong
[3]Curriculum and Instruction Department, International Christian School Hong Kong, Sha Tin, Hong Kong

Correspondence
Ndudi O. Ezeamuzie, Unit 401, Runme Shaw Building, Faculty of Education, University of Hong Kong, Pokfulam Road, Hong Kong.
Email: amuzie@connect.hku.hk

Abstract

Background: Computational thinking is derived from arguments that the underlying practices in computer science augment problem-solving. Most studies investigated computational thinking development as a function of learners' factors, instructional strategies and learning environment. However, the influence of the wider community such as educational policies on computational thinking remains unclear.

Objectives: This study examines the impact of basic and technology-related educational policies on the development of computational thinking.

Methods: Using supervised machine learning, the computational thinking achievements of 31,823 eighth graders across nine countries were analysed. Seven rule-based and tree-based classification models were generated and triangulated to determine how educational policies predicted students' computational thinking.

Results and conclusions: Predictions show that students have a higher propensity to develop computational thinking skills when schools exercise full autonomy in governance and explicitly embed computational thinking in their curriculum. Plans to support students, teachers and schools with technology or introduce 1:1 computing have no discernible predicted influence on students' computational thinking achievement.

Implications: Although predictions deduced from these attributes are not generalizable, traces of how educational policies affect computational thinking exist to articulate more fronts for future research on the influence of educational policies on computational thinking.

KEYWORDS
autonomy, computational thinking, data mining, educational policy, machine learning, policymaking

1 | INTRODUCTION

Wing's (2006) appeal to computer scientists to share the treasures of their domain with everybody represents a seminal juncture in computational thinking (CT) discourse. At the centre of Wing's proposition, lies the claim that inherent practices and concepts in computer science could be harnessed for problem-solving in other domains. Since 'most educators regard problem-solving as the most important learning outcome for life' (Jonassen, 2000, p. 63), the linkage between CT and problem-solving resonated within the educational community and policymakers. For instance, CT has been included as a learning dimension in the national curricula of the European Ministries of Education (Bocconi et al., 2016) and the Next Generation Science Standards in the United States (National Research Council, 2013). Also, considered an essential 21st-century skill, large-scale comparative studies such as the International Computer and Information Literacy Study (ICILS) and Programme for International Student Assessment (PISA) have also incorporated CT in their assessment frameworks (Fraillon et al., 2019; Organisation for Economic Co-operation and Development, 2018).

Besides problem-solving, CT is associated with several educational and social benefits like collaboration, critical thinking, self-management, confidence, mathematical thinking, natural language literacy, reasoning, creativity, metacognition, communication and positive attitudes (Denner et al., 2019; Popat & Starkey, 2019; Scherer et al., 2019). Wing's (2006) viewpoint likened the importance of CT with time-honoured literacies of reading, writing and arithmetic. Though debatable, prospects of CT have necessitated enquiries on the nature and acquisition of CT. Whilst not exhaustive, these include interests in CT's conceptualization (Ezeamuzie & Leung, 2022; Shute et al., 2017), assessment (Korkmaz et al., 2017; Román-González et al., 2017) and development across educational settings like teacher education (Yadav et al., 2018), high schools (Yin et al., 2020), middle schools (Berland & Wilensky, 2015), primary schools (Kong et al., 2018; Pellas, 2023), early childhood (Bers et al., 2014), undergraduates (Ezeamuzie, Leung, Garcia, & Ting, 2022) and special needs (González-González et al., 2019).

Most empirical enquiries on CT focused on instructional strategies (e.g. Hsu et al., 2018; Lye & Koh, 2014; Scherer et al., 2020), learning environments (e.g. Noh & Lee, 2020; Zhang & Nouri, 2019) and other micro-level learners' attributes such as gender, attitude, age and programming experience (e.g. Ezeamuzie, 2023; Sun et al., 2022) and emotions (Pellas, 2023). But, there are wider community contexts such as government educational policies, home environment, school characteristics and classroom settings that may influence students' CT skills (Fraillon et al., 2019). Policies adopted at the national, provincial and school levels play critical roles in shaping educational outcomes. For instance, Hanushek et al. (2013) discovered that granting schools autonomy enhanced the mathematics and science achievements of students in developed economies. How the wider community structures such as the national or local educational policies enhance or inhibit CT development constitutes a critical gap in the CT discourse. To close this gap, this study examines how educational policies could support CT development. Students' CT achievements and educational policy data of the participating nations were captured in the ICILS large-scale comparative study (Fraillon et al., 2019; Fraillon et al., 2020b). The validated big data, which is open and free, contains invaluable data for uncovering the impact of educational policies on the development of CT—a 21st-century problem-solving skill for everybody.

Research Question—*How do the basic and technology-related educational policies predict learners' CT achievement?* The analysis will focus on the following 12 policy features: start age of compulsory education (X1), length of compulsory education (X2), policy direction (X3), autonomy in governance for public schools (X4), autonomy in governance for private schools (X5), curriculum emphasis on CT topics (X6), support for digital learning (X7), mandate for ICT assessment (X8), plans for 1:1 computing @ school (X9), plans to support student with ICT (X10), plans to provide ICT resources (X11) and plan to support teachers with ICT (X12).

2 | BACKGROUND

2.1 | Computational thinking

There is no consensus on both the history and meaning of CT (Ezeamuzie & Leung, 2022). Often, Wing's (2006) view that everybody can harness computer scientists' cognitive styles to solve problems has been cited as a seminal reference (Grover & Pea, 2013; Shute et al., 2017). Although the impact of Wing's call cannot be underestimated, other historical accounts exist (Tedre & Denning, 2016). For instance, CT is also linked to Seymour Papert's work on procedural thinking, which embodies the epistemic development that occurs when children *teach computers to think* (Papert, 1980). Even when not stated explicitly as 'computational thinking', the applicability of computer science practices in problem-solving was engrained in Donald Knuth's view that teaching a computer to perform tasks strengthens conceptual clarity (Knuth, 1974) and the call for the inclusion of programming in liberal arts by Alan Perils, the first recipient of the Association for Computing Machinery Turing Award (Perils, 1962, as cited in Guzdial, 2008).

Wing (2006) associated CT with problem-solving, abstraction, decomposition, system design, parallel processing, recursive thinking, heuristic reasoning and data interpretation. Whereas this framing has been criticized as overly broad (Mannila et al., 2014), several researchers consider Wing's position as a call for further research. Over the years, many frameworks have emerged to articulate the meaning and operationalization of CT. Examples include the CT models by the Computing at School (Csizmadia et al., 2015), the International Society for Technology in Education (Barr et al., 2011; Barr & Stephenson, 2011) and others (e.g. Brennan & Resnick, 2012; Selby & Woollard, 2013; Shute et al., 2017; Weintrop et al., 2016). These frameworks reflect the non-consensus and multifaceted nature of CT, which comprises an array of cognitive styles that reside in computer science.

Though the various models of CT may be encouraging for the nascent field of CT, it is equally important to highlight the inherent challenges in comparing learning and assessment across studies (Ezeamuzie, 2023). Pushing for a consensus model of CT is an uphill if not impossible task. Nonetheless, recent reviews have shown some commonality and convergence in the operationalization. In a systematic review of CT empirical studies, Ezeamuzie and Leung (2022) found that problem-solving and abstraction were consistent constructs in the framing of CT. Although the association between CT and some constructs are established, another tier of concern is the diversity in their operationalization. For example, abstraction takes any of these four processes—discovery, extraction, creation and assembly (Ezeamuzie, 2023; Ezeamuzie, Leung, & Ting, 2022). Other issues of concern include the substantial number of CT studies that had no clear distinction between CT and computer programming

(Ezeamuzie & Leung, 2022; Lye & Koh, 2014). Nardelli (2019) predicted this kind of mix-up will persist if the focus is overly on CT theorization and suggested that CT should be interpreted 'as a shorthand' of computer science for all students (p. 32). In this study, the meaning of CT is interpreted from the definition adopted in the ICILS-2018 large-scale study.

> Computational thinking refers to an individual's ability to recognize aspects of real-world problems which are appropriate for computational formulation and to evaluate and develop algorithmic solutions to those problems so that the solutions could be operationalized with a computer (Fraillon et al., 2019, p. 27).

2.2 | Contextual framework

The underpinning framework in this study derives from the ICILS-2018 contextual framework (Fraillon et al., 2019). Figure 1 shows the relationship between the components of the framework, which encapsulates learning as an outcome of four multilevel contexts: individual, school/classroom, home environment and wider community. Individual dimension captures how students' characteristics and their learning processes impact CT achievement. Classroom/school level entails other factors within the school environment that influence learning, such as learning environment designs and availability of technology. Beyond the school perimeters, the third contextual level deals with the impact of the home environment (e.g. family composition and parental education level) on students' CT achievement. The wider community is the fourth level and encompasses external variables such as a region's technology development index and educational policy.

Although the contextual levels are conceptually demarcated, they are intertwined in practice. For example, technology may exert influence on the school, home, or wider community contextual levels. Within each contextual level, the framework shows how antecedents (features with indirect influence) support the process (features with direct influence) in producing the CT learning outcomes. As captured in the conceptual framework, the antecedent and process are crucial factors of educational policies for delivering outcomes at the wider community level.

2.3 | Educational policies and computational thinking

What constitutes educational policy is broad. In this study, educational policy encompasses the government's directions on the use of ICT for learning, assessment, teacher development and general administration. In addition, four basic characteristics of schooling set by the national or provincial government of a country were examined too: start age of compulsory education, length of compulsory schooling, tier of government that sets policy directions and degree of school autonomy.

Several studies have reported broad differences in the level of ICT access and penetrations across countries (Fraillon et al., 2014; World Bank, 2016). This raises issues about equity and differentials in learning when substantial associations exist between ICT and educational outcomes. For instance, through face-to-face interviews with 36,619 adults in 32 developing countries, the Pew Research Center (2015) found that technology with internet access influences educational outcomes positively. Similarly, a latent analysis of school features across 1727 schools from the ICILS-2013 diet showed that students' digital literacy attainments were distinguishable by the

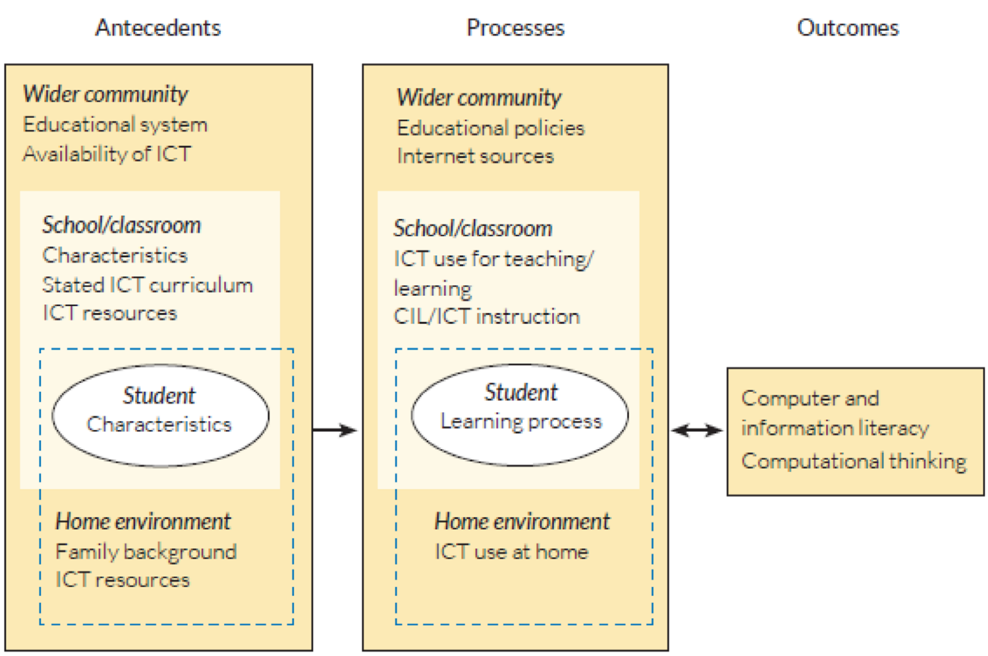

FIGURE 1 Overarching contextual framework (Fraillon et al., 2019).

presence of ICT infrastructure (Gerick, 2018). The above findings are consistent with a second-order meta-analysis ($n = 25$), which revealed the positive impacts of computer technology on students' achievements.

Since computing practices are the underpinnings of CT (Wing, 2006) and most empirical studies operationalized CT obscurely from ICT practices like programming (Ezeamuzie & Leung, 2022), the level of ICT presence and use may influence the development of CT. In fact, ICT in education is increasingly perceived as a standard tradition rather than a peripheral practice (Fraillon et al., 2019) and has a significant correlation with learning outcomes (Gerick et al., 2017). Despite the evidence that ICT affects learners' outcomes and the elevated importance of CT for 21st-century learners, research focusing on CT development is generally limited to personal and micro-level factors like learners' characteristics (e.g. Ezeamuzie, 2023; Sun et al., 2022), learning environments (e.g. Noh & Lee, 2020; Zhang & Nouri, 2019) and instructional strategies (e.g. Hsu et al., 2018; Lye & Koh, 2014; Scherer et al., 2020). ICILS-2018 large-scale study sought to understand how the multilevel contexts affect the development of CT. However, the relationship between educational policies and CT was not examined (Fraillon et al., 2019). On the other hand, it is vivid that policymakers rely on evidence to advance learning and institutional agendas (Wiseman, 2010). Moreover, educators' confidence in the validity of educational practices is enhanced when supported by evidence (Oakley, 2002). The importance of evidence in underpinning policies that guide educational policies cannot be over-emphasized. Therefore, to address the knowledge gap and paucity of investigations into the influence of educational policies on CT, the educational policy attributes captured in ICILS 2018 are analysed in this study.

3 | METHODS

3.1 | Research design and data collection

Data for this study were collected in the ICILS-2018 diet. The multidimensional data can be freely retrieved from the ICILS-2018 database (Mikheeva & Meyer, 2020), which contains the CT and computer and information literacy achievements of 46,561 students in Grade 8 across 14 countries/regions. ICILS is a large-scale comparative study that measures students' preparedness for study, work and life in an information-driven society. ICILS is organized by the International Association for the Evaluation of Educational Achievement (IEA)—an independent non-profit research organization with more than six decades of experience in administering comparative studies including Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS) and International Civic and Citizenship Education Study (ICCS).

Comprehensive details of the ICILS study including the design, sampling, implementation, validation, scaling, data management and creation of the database are documented in the technical report (Fraillon et al., 2020a), database user guide (Mikheeva & Meyer, 2020), assessment framework (Fraillon et al., 2019) and international report (Fraillon et al., 2020b). To summarize, a CT test was administered in 9 out of 14 countries/regions. A total of 31,823 students completed two strands of CT tests, each lasting for 25 minutes with a combined total of 17 questions and 39 score points. The first strand focused on *conceptualizing problems* in an automated bus task, which measured students' knowledge of digital systems, ability to analyse problems and data representation. The second strand focused on *operationalizing solutions* in a farm drone task that assessed students' ability to plan, develop algorithms and evaluate the efficiency of the solutions.

Rasch item response theory model was used to derive students' CT achievement from the 39 score points. The resulting CT scale was standardized with an international mean of 500 and a standard deviation of 100. The CT achievement was based on the scaled difficulties of the assessment items. For each assessment item, ICILS generated a description that articulated the skills and knowledge it measured. An item map was created by ordering the items according to their scaled difficulty—ranging from least difficult to most difficult. With items ordered by difficulty level, an equal number of items were assigned into thirds and categorized into three bands—lower, middle and upper regions. Descriptive interpretations of students' CT achievements were generated from the broad descriptors of the underlying assessment items (see Fraillon et al., 2020a). Students mapped to the *lower* region (less than 459 points) demonstrated an aptitude to use simple sequences in algorithmic problems. However, their solutions contain suboptimal routes in networks and inefficient algorithms in tasks of medium complexity. Students mapped to the *middle* region (459–589 points) understood how CT augments real-world problem-solving with iterative algorithms. Their solutions for medium-complexity tasks are efficient but inefficient for high-complexity problems that require multiple levels of iterations and decisions. Students mapped to the *upper* region (above 589) understood CT as a generalizable problem-solving strategy and designed efficient algorithms for high-complexity tasks.

Educational policy data of the participating countries were collected through an online national context survey addressed to the expert ICILS national research coordinators. Using the IEA IDB Analyser (Mikheeva & Meyer, 2020), the national context data was combined with students' CT achievement. With focus on the basic and ICT-related policies, redundant and unrelated attributes were removed. These included zero-gain information (e.g. unique identifiers of students, schools and countries), students test items, student questionnaire items, variance estimations and sampling weights. Supplementary 1 contains the merged national context and CT achievement data.

Table 1 summarizes the basic and technology-related policy data as collated from the students' CT performance and national context survey in ICILS 2018. The first column of Table 1 shows the resulting 12 educational policy attributes (X1, X2, …, X12) that are potential predictors of CT achievements and their values within the 9 participating countries. Autonomy in governance in public schools (X4), autonomy in governance in private schools (X5), curriculum emphasis on CT

TABLE 1 Summary of basic and technology policy data from ICILS 2018 national context survey and students' computational thinking.

Policy attribute	Value[a]								
Country/region[b]	DEU	DNK	DNW	FIN	FRA	KOR	LUX	PRT	USA
Participants (n = 31,823)	3655 (11.5%)	2404 (7.6%)	1991 (6.3%)	2546 (8.0%)	2940 (9.2%)	2875 (9.0%)	5401 (17.0%)	3221 (10.1%)	6790 (21.3%)
X1: Start Age of Compulsory Education	6	6	6	7	6	7	4	6	6
X2: Length of Compulsory Education	9	10	10	9	10	9	12	12	11
X3: Policy Direction	Provincial	National	Provincial	National	National	National	National	National	Provincial
X4: Autonomy in governance (public schools)	Partial	Partial	Partial	Partial	Partial	Full	Partial	Partial	Partial
X5: Autonomy in governance (private schools)	Full	Full	N/A	Partial	Partial	Full	Full	Full	Full
X6: Curriculum emphasis on CT topics	None	Explicit	None	Implicit	Explicit	Explicit	None	Implicit	Implicit
X7: Support for digital learning	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
X8: Mandate for ICT assessment[c]	No	No	No	No	School	School	No	School	No
X9: Plans for 1:1 computing @ school	Yes	No	Yes	No	No	Yes	Yes	No	No
X10: Plans to support student with ICT	Explicit	Explicit	Explicit	Implicit	Explicit	Explicit	Explicit	Implicit	Implicit
X11: Plans to provide ICT resources	Explicit	Explicit	Explicit	Explicit	Explicit	Explicit	Explicit	Implicit	Explicit
X12: Plan to support teachers with ICT	Explicit	Explicit	Explicit	None	Explicit	Implicit	Explicit	Implicit	Implicit
Y: CT Achievement									
Lower (n = 10,973; 34.5%)	1320 (36.1%)	433 (18.0%)	732 (36.8%)	734 (28.8%)	919 (31.3%)	643 (22.4%)	2578 (47.7%)	1241 (38.5%)	2373 (34.9%)
Middle (*n* = 15,588; 49.0%)	1819 (49.8%)	1408 (58.6%)	993 (49.9%)	1330 (52.2%)	1529 (52.0%)	1326 (46.1%)	2270 (42.0%)	1750 (54.3%)	3163 (46.6%)
Upper (*n* = 5262; 16.5%)	516 (14.1%)	563 (23.4%)	266 (13.4%)	482 (18.9%)	492 (16.7%)	906 (31.5%)	553 (10.2%)	230 (7.1%)	1254 (18.5%)

Note: X1–X12 represent the independent attributes (predictors). Y is the dependent attribute (aka class attribute in supervised learning) and represents the categorical dimension of students CT achievement.

[a]Percentage in the participants' row represents the distribution of students between the 9 participating countries/region. For the row on CT Achievement, the percentage represents the within-country distribution of achievements in 3 grading regions – lower, middle, and upper.

[b]ISO country/region codes: DEU (Germany), DNK (Denmark), DNW (North Rhine-Westphalia, Germany), FIN (Finland), FRA (France), KOR (South Korea), LUX (Luxembourg), PRT (Portugal), USA (United States of America).

[c]Mandate for ICT assessment has no response for options – 'yes, with compulsory assessment', 'yes, using samples', and 'yes, using non-compulsory assessment'.

topics (X6), plans to support students with ICT (X10), plans to provide ICT resources (X11) and plans to support teachers with ICT (X12) were derived as the average of the composite questionnaire items. CT achievement (Y) is the dependent attribute (also known as the class attribute in data mining). Table 2 shows the value–code mapping of the attributes for interpreting the classifiers' rules and trees.

3.2 | Educational data mining approach

The categorical nature of the predictors makes the data mining approach an interpretable model than other statistical methods. Alternative options such as multiple linear regression require continuous variables as predictors. The effectiveness of data mining techniques in knowledge discovery has been reported in several education investigations such as predicting learners' academic performance from their activity in a learning management system (Romero et al., 2013), predicting learners' satisfaction in a massive open online course (Hew et al., 2020) and deconstructing the association between demographic characteristics and learners' academic performance (Rizvi et al., 2019). In data mining, outcomes are sometimes not known a priori but deduced through a set of logically designed processes to uncover patterns between variables. Figure 2 summarizes the three stages of data mining processes in this study. Weka 3.8—

TABLE 2 Code mapping of the attributes' values.

Policy attribute	ICILS code	Value coding (meaning)
X1: Start Age of Compulsory Education	NC2GA02A	inf–5.5 (less than or equal to 5.5 years); 5.5–inf (above 5.5 years)
X2: Length of Compulsory Education	NC2GA02B	inf–10.5 (less than or equal to 10.5 years); 10.5–inf (above 10.5 years)
X3: Overall Policy Direction	NC2GA01	1 (national); 2 (state/provincial)
X4: Autonomy in governance in public school	NC2GA06*A	1 (full autonomy); 2 (partial/some autonomy); 3 (no/little autonomy)
X5: Autonomy in governance in private schools	NC2GA06*B	1 (full autonomy); 2 (partial/some autonomy); 3 (no/little autonomy)
X6: Curriculum emphasis on CT topics	NC2GC20*	1 (explicit); 2 (implicit); 3 (none)
X7: Support for digital learning	NC2GB15	1 (yes); 2 (no)
X8: Mandate for ICT assessment	NC2GB17	4 (yes, but determine at school); 5 (no mandate)
X9: Plans for 1:1 computing @ school	NC2GB14	1 (yes); 2 (no)
X10: Plans to support student with ICT	NC2GB10*	1 (explicit); 2 (implicit); 3 (none)
X11: Plans to provide ICT resources	NC2GB11*	1 (explicit); 2 (implicit); 3 (none)
X12: Plan to support teachers with ICT	NC2GB12*	1 (explicit); 2 (implicit); 3 (none)

Note: *, attributes were derived as the average of the composite questionnaire items. For example, X12 is derived by averaging NC2GB12A, NC2GB12B…, NC2GB12E.

open-source software for machine learning and data mining (Witten et al., 2016), was adopted for data mining operations.

3.2.1 | Instance selection

The class attributes (i.e., CT achievement) has three values—lower, middle and upper. However, the accuracy of classification algorithms is inversely related to the number of options in the class attribute (Witten et al., 2016). Since the central goal of the data mining experiment is to uncover how educational structures may either support or hinder the development of CT skills, retaining the data instances that mapped to 'middle' CT achievement may constitute a noise to the accuracy of the classifiers. Therefore, using data instances that mapped to 'lower' and 'upper' CT achievement was deemed necessary in this analysis.

Using the randomization filter in Weka, the entire data instances were shuffled randomly—a crucial step to prevent selection bias. This dataset was labelled *dataset-3-class* ($n = 31{,}823$) and comprised the data of the entire students mapped to the three CT achievements—lower, middle and upper classes (see Supplementary 2). Another dataset labelled *dataset-2-class* ($n = 16{,}235$) was generated by removing all the instances with middle CT achievement (see Supplementary 3). For comparative insight, the attribute selection and performance of the classifiers were evaluated with both datasets—the ternary class (*dataset-3-class*) and binary class (*dataset-2-class*).

3.2.2 | Attribute selection

The degree of relevance of the attributes constituted another hurdle towards understanding how policies influence students' CT achievement. Pearson Chi-square test of independence was executed and the strength of the association between the 12 predictors and CT achievement was deduced with Cramer's V. Following Romero et al.'s

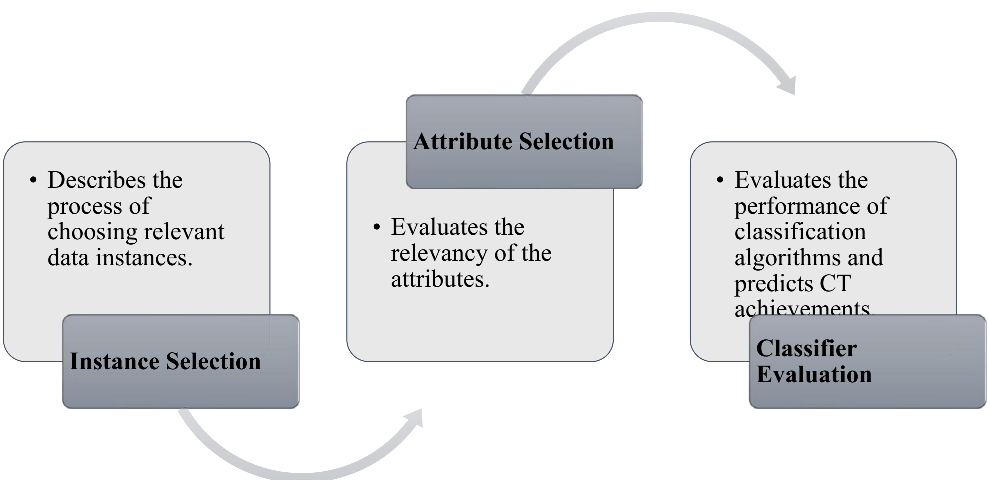

FIGURE 2 Data mining approach for predicting, grouping, and discovering associations in students' computational thinking achievements.

(2013) data mining approach in predicting students' performance through their participation in an online discussion, the relevancy of the attributes was further analysed. Weka provides attribute selection algorithms by (a) ranking attributes based on set filters, (b) searching through the attribute space to select an optimal subset of attributes and (c) wrapping the search and evaluation of sets of attributes by using classification schemes.

In the filter-based ranking, three attribute selection filters that evaluate the relevance of attributes individually were adopted: *ChiSquaredAttributeEval*, *GainRatioAttributeEval* and *InfoGainAttributeEval*. *ChiSquaredAttributeEval* measures Chi-squared statistic as a test of association between predictors and CT class. With respect to the CT class, *GainRatioAttributeEval* and *InfoGainAttributeEval measure* the gain ratio and information gain of the predictors.

– Gain Ratio (Class, Attribute) = (H (Class)—H (Class | Attribute)) / H (Attribute).

– Info Gain (Class, Attribute) = H (Class)—H (Class | Attribute).

– H is the entropy.

Stratified 10-fold validation was implemented for each filter. That is, the dataset was split into 10 equal subsets with each subset retaining the class distribution. The average of the attributes' rank was deducted from the execution of the folds.

The second attribute selection technique entailed executing the *CfsSubsetEval*—a feature selection algorithm that searches for an optimal subset of attributes and evaluates the relevance of subsets of attributes by analysing the predictive capability of individual attributes. Subsets of attributes that are highly correlated with the class but have low intercorrelation are preferred. Thirdly, the wrapper techniques (*WrapperSubsetEval*) evaluated different sets of attributes by using classification schemes. In the evaluation, 10-fold cross-validation was implemented in the wrappers. The following rule-based and tree-based classifiers were wrapped in the evaluation: ZERO, ONER, PART, JRIP, Decision Table, Random Tree, Random Forest, J48, Decision Stump and Hoeffding Tree. The details of the underlying algorithms of the classifiers are described in the Weka guidebook (Witten et al.,2016). In addition, bi-directional or evolutionary searches were adopted in traversing the attribute space. The bi-directional search technique starts at any point, searches in both forward and backward directions and considers all attributes for deletions and additions at all points. Supplementary 4 describes the underpinning classifiers and contains the outputs from the filter-based ranking, wrapper technique and attribute space search.

3.2.3 | Classifier evaluation

The performance of the classification algorithms in predicting students' CT achievement was evaluated. Algorithms learn from the training data and then classifies new observation (test data) into upper or lower regions of CT achievement. In Weka, classification algorithms are broadly categorized into tree-based, rule-based, bayes-based, function-based and lazy algorithms. Tree-based and rule-based classification schemes were adopted because they can be visually interpreted. For each algorithm, the performance was evaluated with the entire data instances (i.e., dataset-3-class) and the data without the 'middle' CT instances (i.e., dataset-2-class). To avoid over-fitting (i.e., when model fits the training data exactly but depreciates in performance with unseen data), pruning was implemented in the algorithms and only rules with more than 5% instance coverage were considered. In addition, stratified 10-fold cross-validation was implemented for each classifier. That is, the dataset was split into 10 equal subsets with each subset retaining the class distribution. For each iteration of the algorithm, a unique subset was used as test data and the remaining 9-folds were used as training data for model fitting.

	Predicted: U	Predicted: L
Actual: U	T_U	F_L
Actual: L	F_U	T_L

FIGURE 3 Confusion matrix of a two-class (upper, lower) of computational thinking achievement. U = Upper, L = Lower, T = True Prediction, F = False Prediction; Accuracy = $(T_U + T_L)/(T_U + F_U + T_L + F_L)$; Precision $_{Upper} = T_U/(T_U + F_U)$; Recall $_{Upper} = T_U/(T_U + F_L)$; Precision $_{Lower} = T_L/(T_L + F_L)$; Recall $_{Lower} = T_L/(T_L + F_U)$.

Accuracy, precision and recall metrics were adopted for evaluating model performances. Figure 3 is the confusion matrix elucidating how the different performance metrics are derived and related. Accuracy represents the percentage of correctly classified instances (i.e., sum of the correct classification / total number of instances in the test). When a class is precited, precision represents the percentage of correct prediction (i.e., number of correct predictions of a class / total number of times the class was predicted). Within an actual class, recall represents the percentage of the class that was correctly predicted (i.e., the number of correct predictions of a class / total number of actual instances in the class).

4 | FINDINGS

4.1 | Summarizing the results

Tying back to the underpinning research question—how the basic and technology-related educational policies predict learners' CT achievement, Table 3 shows the filter-based attribute ranking as the average across the 3 filter selection techniques, the number of times an attribute was nominated in the subset in 10 different seed execution of *CfsSubsetEval* and *WrapperSubsetEval*. Pearson Chi-square test of independence showed statistically significant associations between the 12 predictors and CT achievement at $p = 0.05$. According to Cohen (1988), Cramer's V of 0.1 is classified as small, a value of 0.3 is classified as medium and 0.5 is classified as a large effect. In the ternary-class data, the magnitude of association between the attributes and CT achievement was generally weak except for X1, X2, X4 and X6 which were marginally above the small effect threshold. The

TABLE 3 Attribute selection by average filter ranking, wrapper-based classifier evaluation (WrapperSubsetEval) and optimal subset of attributes (CfsSubsetEval).

Statistics	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Ternary Class CT (lower, middle, upper)												
X^2– Cramer's V	0.132	0.127	0.019	0.134	0.042	0.117	0.080	0.047	0.064	0.019	0.085	0.046
Filters (Average Ranking)	2.4**	4.0**	11.6	2.3**	10.0	1.9**	6.0	9.0	7.5	11.4	4.4**	7.5
CfsSubsetEval (% selected in 10-folds iteration)	100%	80%	–	100%	–	100%	–	–	–	–	80%	–
WrapperSubsetEval (% selected by learning schemes)	80%	–	–	–	–	–	–	–	–	–	–	–
Binary Class CT (lower, upper)												
X^2–Cramer's V	0.154	0.166	0.016	0.181	0.042	0.223	0.106	0.057	0.051	0.022	0.113	0.087
Filters (Average Ranking)	3.3**	3.7**	12	1.7**	10	2.0**	6.0	8.0	9.0	11.0	4.4**	7.0
CfsSubsetEval (% selected in 10-folds iteration)	100%	30%	–	100%	–	100%	–	–	–	–	100%	–
WrapperSubsetEval (% selected by learning schemes)	–	–	–	20%	70%	70%	–	–	–	–	–	–

Note: –, attribute not selected in the subset list of WrapperSubsetEval and CfsSubsetEval.
**Attribute with ranking of relevance in the top half of the attribute list.

TABLE 4 Performance evaluation of the classification algorithms.

			Binary class CT (all attributes)				
				Precision		Recall	
Algorithm	Ternary class CT (all attributes) accuracy	Binary class CT (five high-ranked attributes) accuracy	Accuracy	Lower	Upper	Lower	Upper
ZERO[a]	0.4898	0.6759	0.6759	0.676	–	1.000	0.000
One Rule[a]	0.4995	0.6921	0.6921	0.703	0.585	0.941	0.172
PART[a]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
JRip[a]	0.4898	0.6759	0.7001	0.723	0.577	0.902	0.279
Decision Table[a]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
Random Tree[b]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
Random Forest[b]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
Decision Stump[b]	0.4898	0.6759	0.6759	0.676	–	1.000	0.000
Hoeffding Tree[b]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
J48[b]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279
REP Tree[b]	0.4995	0.6921	0.7001	0.723	0.577	0.902	0.279

Note: The start age of compulsory education (X1), length of compulsory education (X2), autonomy in public school governance (X4), curriculum emphasis on CT (X6) and plan to provide ICT resources (X11) were the 5 high-ranked attributes.
[a]Rule-based.
[b]Tree-based.

association between the attributes and CT achievement in the binary-classed data were similar to the ternary class except for X7 and X11 which were also marginally above the small effect threshold. The filter-based attribute selection revealed important predictors: start age of compulsory education (X1), length of compulsory education (X2), level of autonomy in public school governance (X4), the tangibility of curriculum emphasis on CT (X6) and clarity of plan to provide ICT resources (X11). These attributes also ranked in the top half of the selection algorithms, had stronger associations with CT than other attributes and were selected in *CfsSubsetEval* subsets.

Table 4 shows the performance of the classifiers. The accuracy of classification algorithms was higher when executed with the binary class dataset (i.e., *dataset-2-class* with upper and lower options) than the ternary class equivalent (i.e., *dataset-3-class* with lower, middle and upper). The finding is consistent with the rationale for evaluating the classifiers without the middle CT achievers and aligned to uncover what may support or hinder high CT achievement. The binary-class dataset was further evaluated with the five highest-ranked attributes. However, albeit surprisingly, the accuracy was marginally lower than the performance of classifiers when executed with the all-attribute dataset. Across the classifiers, the precision and recall of the lower CT class are higher than the equivalent metric of the upper CT class. For example, the Decision Table has 72.3% precision and 90.2% recall rates for the lower CT class, which is considerably higher than the 57.7% precision and 27.9% recall rates of the upper CT class. Implicitly, the classification algorithms generated better predictions and sensitivity in identifying lower performers correctly.

4.2 | Interpreting the outputs of the models

To interpret the outputs, 70% accuracy was designated as the cut-off for selecting classifiers. Figure 4 through Figure 10 show the generated rules and trees. Although the outcomes of the classifiers are represented with different notations, they can be reduced to IF-ELSE-THEN rules.

In the Decision Table classification (Figure 4), each row represents one of the seven rules. The 7th rule, for example, is interpreted as '*IF* X4 = 1 AND X5 = 1 AND X6 = 1 THEN CT = upper'. When interpreted with code-value mapping (see Table 2), it denotes that within an accuracy of 70%, students are predicted to attain upper CT level when educational policy support *full* autonomy in public school governance (X4), *full* autonomy in private school governance (X5) and follow curriculum with explicit emphasis on CT topics (X6). The other 6 rules are interpreted in a similar approach. In fact, the 2nd rule is the inverse of the 7th rule, showing the predicted likelihood of lower CT achievement when schools have only partial autonomy and emphasis on CT in the curriculum is only implied.

JRip classification (Figure 5) yielded 2 rules with the '=>' representing the THEN operator. The first rule states that 'IF X6 = 1 AND X5 = 1 THEN CT = upper'. Implicitly, with an accuracy of 70%, students will attain upper CT levels when the educational policy implements a curriculum with explicit emphasis on CT topics (X6) and support *full* autonomy in private school governance (X5). The second rule denotes that if the other conditions are not satisfied, then the students are predicted to attain lower CT achievement. At the end of each rule is a bracket with two numbers separated by '/'. The first number shows coverage (i.e., the number of instances evaluated with the rule) and the second number represents the number of incorrectly classified instances. The second number / first number shows the classification error of a particular rule. For instance, the first rule has a classification error of 42.3%.

PART classification (Figure 6) generated 4 rules with ':' representing the THEN operator. According to the generated rules, students will attain lower CT levels when the educational policy has the following characteristics: public schools have only *partial* autonomy in governance (X4), the start age for compulsory education is above 5.5 years (X1), no explicit plans to provide ICT resources (X11) and overarching education direction is based on state/provincial government. Similarly, prediction shows a propensity for lower CT achievements when the duration of compulsory education is greater than 10.5 (X2) and private schools have only *partial* autonomy in governance (X5).

Some tree-based classification algorithms also met the 70% accuracy threshold: Hoeffding Tree (Figure 7), J48 (Figure 8), Random Tree (Figure 9) and REP Tree (Figure 10). These trees comprise nodes (the attributes) and branches (the value of the attributes). On the top of the figures is the root node, which branches to the leaf nodes. For example, the root node in Hoeffding Tree (Figure 7) is Curriculum emphasis on CT topics (X6) and has three branches: 1 (explicit), 2 (implicit) and 3 (none), representing the levels of emphasis on CT. Each path from the root terminates in a decision node, which has a CT class label of either lower or upper. The rules are generated by tracing the paths from the root to the decision nodes. For instance, tracing the path from the left in Hoeffding Tree (Figure 7), the first rule predicts that when an educational policy follows a curriculum with *explicit* (1) emphasis on CT topics (X6) and support *full (1)* autonomy in private school governance (X5), then the students will probably attain upper CT level.

```
Rules:
==================================
X4    X5    X6    CT
==================================
2     9     3     lower
2     1     3     lower
2     2     2     lower
2     1     2     lower
2     2     1     lower
2     1     1     upper
1     1     1     upper
==================================
```

FIGURE 4 Rules generated by Decision Table classification algorithm.

```
JRIP rules:
===========

(X6 = 1) and (X5 = 1) => CT=upper (2545.0/1076.0)
 => CT=lower (13690.0/3793.0)
```

FIGURE 5 Rules generated by JRIP classification algorithm.

```
PART decision list
------------------

X4 = 2 AND
X1 = '(5.5-inf)' AND
X11 = 1 AND
X3 = 2: lower (6461.0/2036.0)

X2 = '(10.5-inf)': lower (4602.0/783.0)

X5 = 2: lower (2627.0/974.0)

: upper (2545.0/1076.0)
```

FIGURE 6 Rules generated by PART classification algorithm.

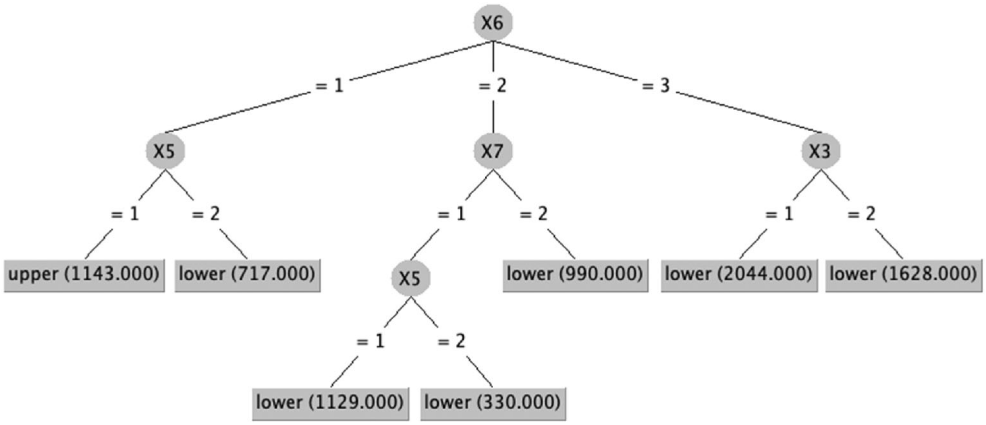

FIGURE 7 Tree generated by Hoeffding Tree classification algorithm.

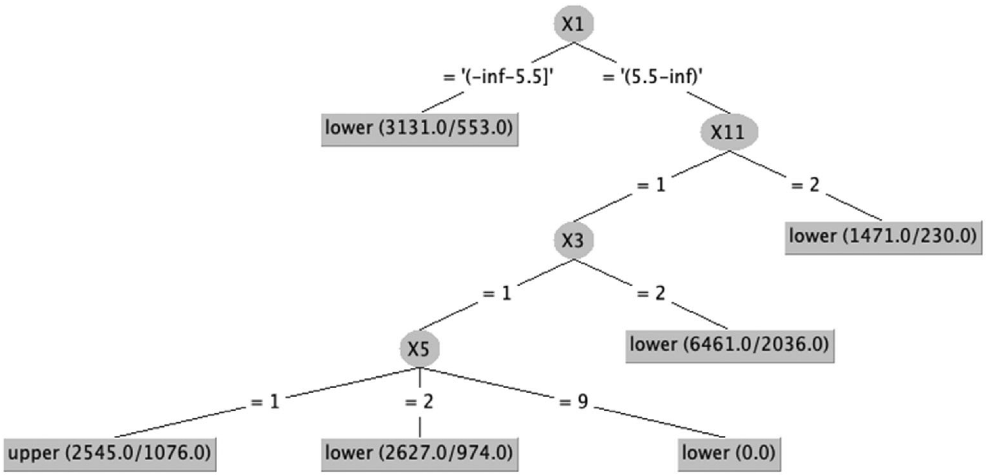

FIGURE 8 Tree generated by J48 classification algorithm.

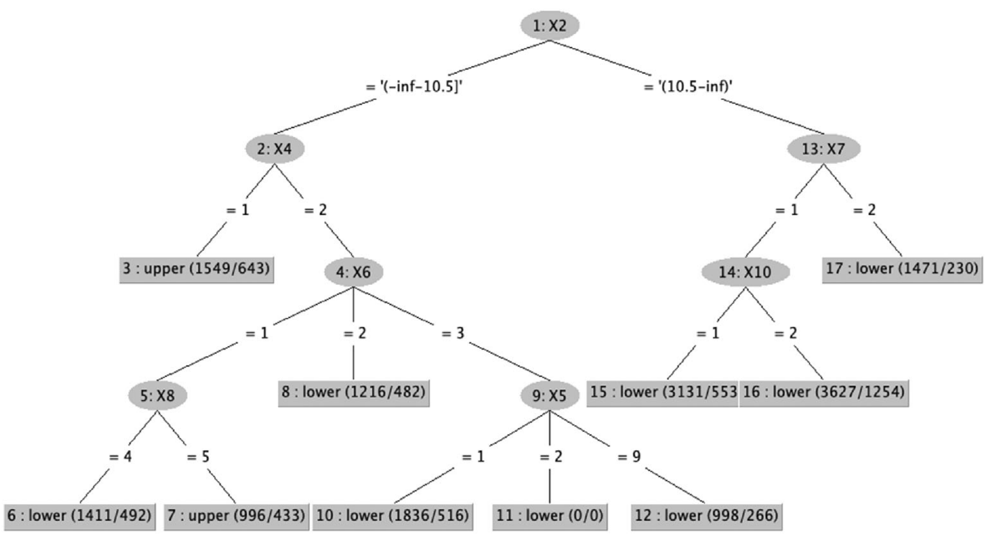

FIGURE 9 Tree generated by Random Tree classification algorithm.

5 | DISCUSSION

The generated rules from the rule-based and tree-based classifiers provided human-readable explanations. Relying on rules from one classifier does not offer coverage for the attributes and is prone to bias. To predict the impacts of the attributes in promoting or undermining CT achievement, the rules generated by seven classifiers are triangulated.

Start age of compulsory education (X1)—although selected as a relevant attribute, the influence on students' CT achievement from the classifiers is contradictory. In PART (Figure 6), start age for compulsory education that is greater than 5.5 years is a component of a rule

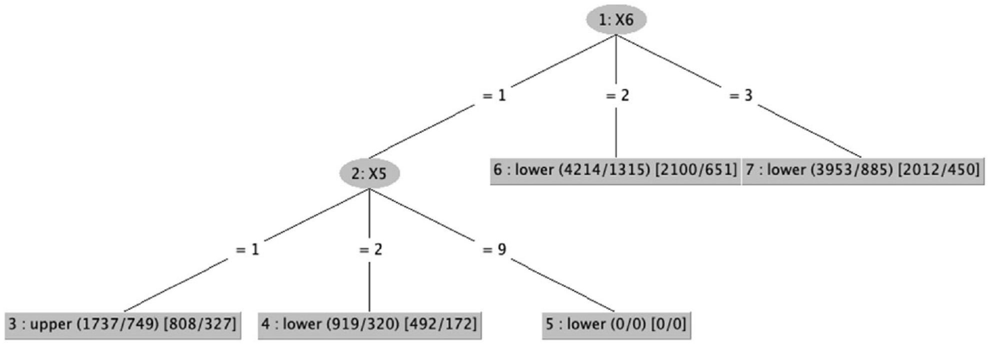

FIGURE 10 Tree generated by REP Tree classification algorithm.

that predicted lower CT achievement. This contradicts a rule in J48 classifier (Figure 8) that predicted lower CT achievement when the start age for compulsory education is less than 5.5 years.

Length of compulsory education (X2)—similar to the start age of compulsory education, this attribute was selected as a relevant attribute for predicting CT achievement. PART (Figure 6) predicted lower CT achievement when the duration of compulsory education is greater than 10.5 years. This is consistent with a rule in Random Tree (Figure 9) that predicts upper CT achievement when the duration of compulsory education is less than 10.5 years and public schools have full autonomy in governance (X4). Ordinarily, a longer education duration will be expected to yield upper CT achievement, which contradicts the present finding. However, participants in the ICILS study were recruited from Grade 8 (i.e., after 8 years of compulsory education), which cast doubt over the applicability of the rules especially when the duration of compulsory education is greater than 10.5 years.

Overarching direction of education (X3) is ranked least relevant by the attribute selection algorithms. Whether the national or provincial government has overall oversight, the evidence about its predicting capability on CT development appears weak. Appearing only as a component of the rules in PART (Figure 6) and J48 (Figure 8), lower CT achievement was predicted when the educational direction is set at the provincial/state level. However, when there is no emphasis on CT topics in the Hoeffding Tree (Figure 7), lower CT was predicted notwithstanding whether the overall direction of education was set at national or provincial levels.

Autonomy in governance in *public and private schools (X4, X5)*—according to the classifiers, students are more likely to attain upper CT when schools exercise full autonomy in the overall governance. Similarly, a lower CT achievement is predicted for schools that exercise little or partial autonomy in governance. For instance, the Decision Table (Figure 4) shows that students will attain upper CT levels when educational policy support *full* autonomy in public (X4) and private (X5) school governance. Inversely, the Decision Table predicted lower CT achievement when both public and private schools have only partial autonomy in governance. These findings are consistent across the classifiers, appearing in the school autonomy's branches leading to the decision nodes. For example, when the educational policy implements a curriculum with explicit emphasis on CT, the Hoeffding Tree (Figure 7) predicts upper and lower CT achievements for students in private schools that support *full and partial* autonomy in governance, respectively.

Curriculum emphasis on CT (X6) was selected as a relevant attribute for predicting CT achievement. The rules show that highlighting CT explicitly in the curriculum was associated with upper CT achievement. JRip (Figure 5), the Hoeffding Tree (Figure 7) and REP Tress (Figure 10) predicted that students will attain upper CT levels when educational policy implements a curriculum with explicit emphasis on CT, especially when private schools have full autonomy in governance. Whenever implicit or no emphasis on CT is made in a curriculum, the classifiers predicted lower CT achievement. One possible explanation that subsists is that explicit integration into the curriculum adds clarity to CT literacy and prompts educators to query how the educational goals could be achieved.

Support for digital learning (X7)—no inference could be deduced from either the presence or absence of support for digital learning on students' CT performance. Hoeffding Tree (Figure 7) and Random Tree (Figure 9) evaluated the influence of support for digital learning as a node in the rules. Irrespective of the branching of 'yes' or 'no' in the rules, the algorithm predicted lower CT achievement.

Mandate for ICT assessment (X8)—evidence about the predicting capability of this attribute on CT development is weak. Random Tree (Figure 9) contains the only reference to the mandate for ICT assessment, appearing as the decision node in a 4-node rule. The classifier mapped the use of school-based assessment and the absence of mandated assessment to lower and upper CT achievements, respectively. However, considering the antecedents of this decision, including the requirement for partial autonomy in public school governance and non-evaluation in other classifiers, the prediction appears rather weak.

Plans for 1:1 computing (X9)—no conclusion could be deduced from either the presence or absence of plans for 1:1 computing on students' CT performance. In addition, the attribute selection algorithm flagged this variable as one but least in relevancy.

Plans to support students with ICT (X10)—No prediction could be deduced from the nature of the plans to support students with ICT on students' CT achievement. Random Tree (Figure 9) evaluated the influence of plans to support students with ICT as a node in the rules. Yet, irrespective of the branching of 'explicit' or 'implicit' in the rules, the algorithm predicted lower CT achievement.

Plans to provide ICT resources (X11)—Although this feature was selected by the attribute selection algorithms, the evidence about the predicting capability on CT development appears weak. There is no direct inference about this attribute and appeared only as a component of rules in PART (Figure 6) and J48 (Figure 8). In PART's rule, lower CT achievement was predicted when there is an explicit plan to provide ICT resources. On the contrary, the rule set in J48 predicted similar lower CT achievement but for an implicit plan to provide ICT resources.

Plans to support teachers with ICT (X12) did not appear in any of the rules generated in the 7 classification algorithms. Hence, no conclusion could be deduced from whether the plans to support teachers with ICT are explicit or implicit.

6 | CONCLUSION, LIMITATIONS AND FUTURE RESEARCH

With the increasing emphasis on CT as a component of 21st century skills, extant empirical investigations have focused on the influence of learners' characteristics, instructional strategies and learning environments. Although it is obvious that government educational policies influence educational outcomes, policymakers are equally advised to rely on evidence to advance learning and institutional agendas (Wiseman, 2010). How government policies, investments and educational direction may influence students' CT development remains unknown.

To highlight the contributions of this work as encapsulated in the research question—how the basic and technology-related educational policies predict learners' CT achievement, it is essential to acknowledge the limitations. Findings from this study were based on an accuracy cut-off of 70%. Although no fixed rule exists for an acceptable accuracy level in data mining, higher accuracy is desirable and enhances confidence in the classifiers. Nonetheless, the cut-off accuracy is better than random probability guesses between upper and lower (50%) and the accuracy of some basic classifiers such as ZERO (67.6%), which predicts the class with the highest count (Witten et al., 2016). Also, the recall of the upper CT class in the classifiers was less than 30%, which implies that models have less sensitivity in predicting high CT achievers. Another threat to the validity of the finding is inherited from the data. In predicting CT achievement, every region's current policy was associated with students' performance. However, this does not account for changes that might have occurred over the years and before the participants attained Grade 8. For example, a region might have adjusted compulsory education's start age or length. The non-consensus operationalization of CT across schools and countries constitutes another element of bias. Whereas the assessment of CT in ICILS was based on the adopted definition for this study (see Computational Thinking section), CT is still an emerging educational construct with varied meanings (Ezeamuzie & Leung, 2022). Given these limitations, the findings from this study may not be generalized. This study more appropriately sets the ball rolling and extends our knowledge towards articulating how policies may affect learners' CT achievement. Hence, while the major discoveries from this study are highlighted below, it is equally important to note that the predicting patterns may change as CT crystallizes.

Given the limitations, one insight from this study is that students may have a higher propensity to attain upper CT levels when the educational policy supports schools to exercise full autonomy in governance and have CT explicitly embedded in the curriculum. This is consistent with the findings in Hanushek et al.'s (2013) cross-country panel analysis of four waves of the triennial Programme for International Student Assessment (PISA) that assessed the mathematics and science achievements of more than one million 15-year-old students in 42 countries. Hanushek et al. (2013) discovered that autonomy enhanced the achievement of learners in developed economies. Since CT is still emerging conceptually, how schools may exercise autonomy seems unclear. Plausibly, schools can adapt, interpret and specify CT within their respective contextual settings. In addition, the positive impact of CT explicitness in the curriculum is aligned with the recommendation for CT practices to be made explicit in learning as they are often latent in everyday problem-solving (Ezeamuzie, Leung, Garcia, & Ting, 2022).

Moreso, another prediction from this study shows that the start age of compulsory education, length of compulsory education, the body responsible for the overarching direction of education, support for digital learning and mandates for ICT assessment have conflicting findings and are weak predictors of CT achievement. Four attributes expressed only plans for changes in the educational policy: plans for 1:1 computing, plans to support students with ICT, plans to provide ICT resources and plans to support teachers with ICT. No predictions were deduced from attributes that expressed only plans. As noted in the limitations, the weak or no predictions deduced from these attributes may not imply that the attributes are not relevant. Rather, it opens more fronts for future research and discussion on the influence of educational policies on CT. Specifically, future studies should examine the impact of real and executable policies instead of just the plans.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/jcal.13041.

DATA AVAILABILITY STATEMENT

The data supporting this study's findings are available for free and can be accessed from the ICILS-2018 database (Mikheeva & Meyer, 2020).

RESEARCH INVOLVING HUMAN PARTICIPANTS AND/OR ANIMALS

This research involves the re-use of data that are already in the public domain. Approval for the study and preparation of the open database of the large-scale comparative was obtained from the ethics committee of The International Association for the Evaluation of Educational Achievement (IEA) and the Australian Council for Educational

Research. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

ORCID

Ndudi O. Ezeamuzie https://orcid.org/0000-0001-8946-5709
Jessica S. C. Leung https://orcid.org/0000-0002-6299-8158
Dennis C. L. Fung https://orcid.org/0000-0002-1844-5468
Mercy N. Ezeamuzie https://orcid.org/0000-0001-5966-7772

REFERENCES

Barr, D., Harrison, J., & Conery, L. (2011). Computational thinking: A digital age skill for everyone. *Learning Leading with Technology*, *38*(6), 20–23. https://www.learningandleading-digital.com/learning_leading/20110304?pm=2&pg=22#pg22

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community? *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

Berland, M., & Wilensky, U. (2015). Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. *Journal of Science Education and Technology*, *24*(5), 628–647. https://doi.org/10.1007/s10956-015-9552-x

Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers & Education*, *72*(C), 145–157. https://doi.org/10.1016/j.compedu.2013.10.020

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., & Engelhardt, K. (2016). *Developing computational thinking in compulsory education - implications for policy and practice*. Publications Office of the European Union. https://doi.org/10.2791/792158

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American Educational Research Association* (Vol. 1, pp. 1–25). http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Csizmadia, A., Curzon, P., Dorling, M., Humphreys, S., Ng, T., Selby, C., & Woollard, J. (2015). Computational thinking: A guide for teachers. https://eprints.soton.ac.uk/424545/

Denner, J., Campe, S., & Werner, L. (2019). Does computer game design and programming benefit children? A meta-synthesis of research. *ACM Transactions on Computing Education*, *19*(3), Article 19. https://doi.org/10.1145/3277565

Ezeamuzie, N. O. (2023). Abstractive-based programming approach to computational thinking: Discover, extract, create, and assemble. *Journal of Educational Computing Research*, *61*(3), 605–638. https://doi.org/10.1177/07356331221134423

Ezeamuzie, N. O., & Leung, J. S. C. (2022). Computational thinking through an empirical lens: A systematic review of literature. *Journal of Educational Computing Research*, *60*(2), 481–511. https://doi.org/10.1177/07356331211033158

Ezeamuzie, N. O., Leung, J. S. C., Garcia, R., & Ting, F. S. T. (2022). Discovering computational thinking in everyday problem solving: A multiple case study of route planning. *Journal of Computer Assisted Learning*, *38*(6), 1779–1796. https://doi.org/10.1111/jcal.12720

Ezeamuzie, N. O., Leung, J. S. C., & Ting, F. S. T. (2022). Unleashing the potential of abstraction from cloud of computational thinking: A systematic review of literature. *Journal of Educational Computing Research*, *60*(4), 877–905. https://doi.org/10.1177/07356331211055379

Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., & Friedman, T. (2019). *IEA international computer and information literacy study 2018: Assessment framework*. Springer. https://doi.org/10.1007/978-3-030-19389-8

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2020a). *IEA international computer and information literacy study 2018: Technical report*. IEA. https://www.iea.nl/publications/technical-reports/icils-2018-technical-report

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2020b). *Preparing for life in a digital world: IEA international computer and information literacy study 2018 international report*. Springer. https://doi.org/10.1007/978-3-030-38781-5

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age: The IEA international computer and information literacy study international report*. Springer. https://doi.org/10.1007/978-3-319-14222-7_1

Gerick, J. (2018). School level characteristics and students' CIL in Europe – A latent class analysis approach. *Computers & Education*, *120*, 160–171. https://doi.org/10.1016/j.compedu.2018.01.013

Gerick, J., Eickelmann, B., & Bos, W. (2017). School-level predictors for the use of ICT in schools and students' CIL in international comparison. *Large-Scale Assessments in Education*, *5*(1), Article 5. https://doi.org/10.1186/s40536-017-0037-7

González-González, C. S., Herrera-González, E., Moreno-Ruiz, L., Reyes-Alonso, N., Hernández-Morales, S., Guzmán-Franco, M. D., & Infante-Moro, A. (2019). Computational thinking and down syndrome: An exploratory study using the Kibo robot. *Informatics*, *6*(2), Article 25. https://doi.org/10.3390/informatics6020025

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189x12463051

Guzdial, M. (2008). Paving the way for computational thinking. *Communications of the ACM*, *51*(8), 25–27. https://doi.org/10.1145/1378704.1378713

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212–232. https://doi.org/10.1016/j.jdeveco.2012.08.002

Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education*, *145*, Article 103724. https://doi.org/10.1016/j.compedu.2019.103724

Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education*, *126*, 296–310. https://doi.org/10.1016/j.compedu.2018.07.004

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85. https://doi.org/10.1007/bf02300500

Knuth, D. E. (1974). Computer science and its relation to mathematics. *The American Mathematical Monthly*, *81*(4), 323–343. https://doi.org/10.1080/00029890.1974.11993556

Kong, S.-C., Chiu, M. M., & Lai, M. (2018). A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education. *Computers & Education*, *127*, 178–189. https://doi.org/10.1016/j.compedu.2018.08.026

Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior*, *72*, 558–569. https://doi.org/10.1016/j.chb.2017.01.005

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior*, *41*, 51–61. https://doi.org/10.1016/j.chb.2014.09.012

Mannila, L., Dagiene, V., Demo, B., Grgurina, N., Mirolo, C., Rolandsson, L., & Settle, A. (2014). Computational thinking in K-9 education. In A. Clear & R. Lister (Eds.), *Proceedings of the working group reports of the 2014 on Innovation & Technology in computer science education conference* (pp. 1–29). ACM. https://doi.org/10.1145/2713609.2713610

Mikheeva, E., & Meyer, S. (2020). *IEA international computer and information literacy study 2018: User guide for the international database*. IEA. https://www.iea.nl/publications/user-guides/icils-2018-user-guide-international-database

Nardelli, E. (2019). Do we really need computational thinking? *Communications of the ACM*, *62*(2), 32–35. https://doi.org/10.1145/3231587

National Research Council. (2013). *Next generation science standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Noh, J., & Lee, J. (2020). Effects of robotics programming on the computational thinking and creativity of elementary school students. *Educational Technology Research and Development*, *68*(1), 463–484. https://doi.org/10.1007/s11423-019-09708-w

Oakley, A. (2002). Social science and evidence-based everything: The case of education. *Educational Review*, *54*(3), 277–286. https://doi.org/10.1080/0013191022000016329

Organisation for Economic Co-operation and Development. (2018). *PISA 2021 mathematics framework (draft)*. OECD Publishing. https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Pellas, N. (2023). Exploring relationships among students' computational thinking skills, emotions, and cognitive load using simulation games in primary education. *Journal of Computer Assisted Learning*, *39*, 1576–1590. https://doi.org/10.1111/jcal.12819

Pew Research Center. (2015). *Internet seen as positive influence on education but negative on morality in emerging and developing nations*. Pew Research Center. https://assets.pewresearch.org/wp-content/uploads/sites/2/2015/03/Pew-Research-Center-Technology-Report-FINAL-March-19-20151.pdf

Popat, S., & Starkey, L. (2019). Learning to code or coding to learn? A systematic review. *Computers & Education*, *128*, 365–376. https://doi.org/10.1016/j.compedu.2018.10.005

Rizvi, S., Rienties, B., & Khoja, S. A. (2019). The role of demographics in online learning; a decision tree based approach. *Computers & Education*, *137*, 32–47. https://doi.org/10.1016/j.compedu.2019.04.001

Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the computational thinking test. *Computers in Human Behavior*, *72*, 678–691. https://doi.org/10.1016/j.chb.2016.08.047

Romero, C., López, M.-I., Luna, J.-M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*, *68*, 458–472. https://doi.org/10.1016/j.compedu.2013.06.009

Scherer, R., Siddiq, F., & Viveros, B. S. (2019). The cognitive benefits of learning computer programming: A meta-analysis of transfer effects. *Journal of Educational Psychology*, *111*(5), 764–792. https://doi.org/10.1037/edu0000314

Scherer, R., Siddiq, F., & Viveros, B. S. (2020). A meta-analysis of teaching and learning computer programming: Effective instructional approaches and conditions. *Computers in Human Behavior*, *109*, Article 106349. https://doi.org/10.1016/j.chb.2020.106349

Selby, C., & Woollard, J. (2013). Computational thinking: the developing definition. https://eprints.soton.ac.uk/356481/

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, *22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003

Sun, L., Hu, L., & Zhou, D. (2022). Programming attitudes predict computational thinking: Analysis of differences in gender and programming experience. *Computers & Education*, *181*, Article 104457. https://doi.org/10.1016/j.compedu.2022.104457

Tedre, M., & Denning, P. J. (2016). The long quest for computational thinking. In J. Sheard & C. S. Montero (Eds.), *Proceedings of the 16th Koli calling international conference on computing education research* (pp. 120–129). ACM. https://doi.org/10.1145/2999541.2999542

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, *25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5

Wing, J. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Wiseman, A. W. (2010). The uses of evidence for educational policymaking: Global contexts and international trends. *Review of Research in Education*, *34*(1), 1–24. https://doi.org/10.3102/0091732x09350472

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

World Bank. (2016). *World development report 2016: Digital dividends*. World Bank Publications. https://doi.org/10.1596/978-1-4648-0671-1

Yadav, A., Krist, C., Good, J., & Caeli, E. N. (2018). Computational thinking in elementary classrooms: Measuring teacher understanding of computational ideas for teaching science. *Computer Science Education*, *28*(4), 371–400. https://doi.org/10.1080/08993408.2018.1560550

Yin, Y., Hadad, R., Tang, X., & Lin, Q. (2020). Improving and assessing computational thinking in maker activities: The integration with physics and engineering learning. *Journal of Science Education and Technology*, *29*(2), 189–214. https://doi.org/10.1007/s10956-019-09794-8

Zhang, L., & Nouri, J. (2019). A systematic review of learning computational thinking through scratch in K-9. *Computers & Education*, *141*, Article 103607. https://doi.org/10.1016/j.compedu.2019.103607

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ezeamuzie, N. O., Leung, J. S. C., Fung, D. C. L., & Ezeamuzie, M. N. (2024). Educational policy as predictor of computational thinking: A supervised machine learning approach. *Journal of Computer Assisted Learning*, 1–14. https://doi.org/10.1111/jcal.13041